Univerzita Karlova v Praze
Přírodovědecká fakulta

Studijní program: Biologie

Studijní obor: Genetika, molekulární biologie a virologie

Ing. Anastasiya Pitkina

Mikrobiální společenstva a metagenom průmyslově znečištěných půd: výskyt genů kódujících AEH

Microbial consortia and metagenome of industrially polluted soil: occurrence of genes encoding AEH

Diplomová práce

Vedoucí diplomové práce / Školitel:
RNDr. Pavel Kyslík, CSc.

Prague, 2015

Svoluji k zapůjčení své diplomové práce ke studijním účelům a prosím, aby byla vedena přesná evidence vypůjčovatelů. Převzaté údaje je vypůjčovatel povinen řádně ocitovat.

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 4.05.2015

Podpis:

Anastasiya Pitkina

Poděkování:

Chtěla bych poděkovat celému kolektivu Laboratoře enzymových technologií Mikrobiologického ústavu, v.v.i AV ČR za poskytnutou odbornou pomoc a přátelskou atmosféru po celou dobu praxe v laboratoři. Speciální poděkovaní patří vedoucímu této diplomové práce panu RNDr. Pavlu Kyslíkovi, CSc., za cenné rady a pomoc při jejím psaní; dále Ing. Evě Kyslíkové, CSc. a RNDr. Andree Palyzové, PhD za pomoc při kultivaci mikroorganizmů včetně přítokované kultivace; RNDr, Heleně Marešové, CSc., Mgr. Jiřímu Zahradníkovi, RNDr. Václavu Štěpánkovi, CSc. a zejména RNDr. Martině Plačkové za pomoc při plánování molekulárně genetických pokusů. V neposlední řadě také děkuji kolektivu Laboratoře environmentální mikrobiologie pod vedením RNDr. Petra Baldriana, PhD za spolupráci na zkoumání mikrobiální diverzity metodou Next-generation sequencing.

Tato práce byla vypracována v letech 2013-2015 s finanční podporou Grantové agentury Univerzity Karlovy GAUK č. 243-259544 a výzkumného a vývojového projektu RVO 61388971 MBÚ AV ČR, v.v.i.

Abstrakt

Půda obsahuje různorodá společenstva bakterii, což z ní děla atraktivní cíl pro metagenomický výzkum a zkoumání kultivovatelných organizmů. Tato diplomová práce analyzuje složení mikrobiálního společenstva farmaceuticky znečištěných půd, s využitím sekvenační technologie nové generace Illumina a knihovny amplikonů 16S rDNA. Tato analýza odhalila vysokou komplexitu mikrobiálního prostředí půd a potvrdila to, že antropogenní aktivita (konkrétně produkce beta-laktamových antibiotik) ovlivňuje variabilitu a relativní zastoupení druhů, aniž by však snižovala mikrobiální diverzitu. V druhé části této diplomové práce je popsaná isolace a heterologní exprese nového genu kódujícího alfa-amino acid ester hydrolázu (AEH), který pochází z kultivovatelného půdního mikroorganizmu *B. cereus.* AEH mají velký průmyslový potenciál v biokatalytických syntézách beta-laktamových antibiotik, které jsou momentálně velmi důležité z hlediska uplatnění v medicíně.

Klíčová slova

16S rRNA, AEH, alfa-amino acid ester hydroláza, *B. cereus*, Illumina, metagenomika, molekulární klonování

Abstract

Soils contain highly diverse consortia of bacteria making them very attractive starting points for both culture-dependent and metagenomic discovery efforts. The present diploma thesis analyses the composition of the microbial community from pharmaceutically polluted soil, with the employment of next-generation Illumina sequencing of 16S rDNA region. This analysis revealed high complexity of the soil microbial environment and confirmed that anthropogenic activity (represented by production of beta-lactam antibiotics) influences the variability and abundance of the species, yet without reducing the microbial diversity. In the second part of the thesis, isolation and heterologous expression of a novel gene encoding alpha-amino acid ester hydrolase (AEH) from a cultivable soil microorganism *B. cereus* is described. AEHs possess industrial potential for biocatalytic synthesis of semi-synthetic beta-lactam antibiotics, which are presently of great clinical importance.

Key words

16S rRNA, AEH, alpha-amino acid ester hydrolase, *B. cereus*, Illumina, metagenomics, molecular cloning

Table of Contents

List of Abbreviations

6-APA	6-aminopenicillanic acid
7-ACA	7-aminocephalosporanic acid
7-ADCA	7-aminodeacetoxycephalosporanic acid
ACV	L-delta-(alpha-aminoadipoyl)-L-cysteinyl-D-valine
AEH	Alpha-amino acid ester hydrolase
ANOVA	Analysis of variance
CDW	Cell dry weight
CODEHOP	COnsensus- DEgenerate Hybrid Oligonucleotide Primer
CTAB	Cetyl trimethyl ammonium bromide
DAAO	D-amino acid oxidase
DAB	Dimethyl amino benzaldehyde
DAC	Deacetylcefalosporin C
DACS	Deacetylcefalosporin C synthase
DAOC	Deacetoxycefalosporin C
DAOCS	Deacetoxycefalosporin C synthase
ddNTPs	Dideoxynucleotide triphosphates
DMSO	Dimethyl sulfoxide
dNTP	Deoxynucleoside triphosphate
ENS	Effective number of species
GAC	Glutaryl acylase
IPNS	Isopenicillin N synthase
MANOVA	Multivariate analysis of variance
NGS	Next-Generation Sequencing
OTU	Operational taxonomic unit
PEG	Polyethylene glycol
PGA	Penicillin G acylase
PGME	Phenylglycine methyl ester
SA	Specific activity
SSBA	Semi-synthetic beta-lactam antibiotic
VA	Volumetric activity
WCB	Working cell bank

Introduction

Soil is a highly complex system containing a wide variety of bacterial communities. Soils from diverse locations around the world have been explored to develop new medicinal and industrial applications. The study of community changes in respect to beta-lactam antibiotic presence, as a type of anthropological pressure, can provide an early warning system for future industrially and clinically relevant investigations; among other helping develop next-generation antibiotics active against resistant strains.

This diploma thesis consists of two thematic parts. The first part, entitled microbial consortia of industrially polluted soils, focuses on acquiring knowledge about communities of microorganisms, influenced by anthropogenic activity, namely by the presence of antibiotics (mostly beta-lactams) in the environment. Bacterial community dwelling in the immediate vicinity to a pharmaceutical plant will be characterized as to its composition and compared to a community from unpolluted soil by Illumina sequencing of the V4 region of 16S rDNA.

Soil microorganisms have been the most valuable source of useful natural products, including industrially important antibiotics and biocatalysts. It is most likely the ecosystem where antibiotic synthesis originally evolved (D'Costa et al., 2007). It is therefore assumed that microbial consortia from pharmaceutically polluted soils can harbour increased diversity and heterogeneity of antibiotic metabolizing enzymes (Torres-Cortés et al., 2011), potentially including alpha-amino ester hydrolases (AEH), which will be in the spotlight of the second part of the present thesis.

To obtain medically exploitable beta-lactams either chemical synthesis or biocatalysis can be applied. Alpha-amino ester hydrolases have been a subject of close attention recently, due to their potential as biocatalysts in semi-synthetic beta-lactam antibiotic (SSBA) biotransformation. At present, a strong trend can be observed towards sustainable biotechnological production, providing products of high quality with lower environmental burden and pursuing goals related to the so-called "green chemistry" (Kumar and Prasad, 2011). Beta-lactam antibiotics are a good illustration of this trend, because their production has evolved from almost completely dependent on chemical syntheses towards the use of biocatalytic processes with substantial environmental and economic benefits (Thykaer and Nielsen, 2003).

Much effort is currently invested into developing and improving enzyme-based methods of synthesizing beta-lactams. There are several examples of successful industrial application of enzymatic production of beta-lactams, e.g. DSM company (van Langen et al.), with penicillin G acylase (PGA) remaining the most widely applied biocatalyst. When compared to PGA, AEH offers a number of advantages, but the information available on its biochemistry is currently not sufficient, and the potential of AEH as a biocatalyst has not been fully investigated. Since only ten strains have been described to produce AEHs to date, the present thesis is aimed to contribute to this exploration.

1 Literature review

1.1 Microbial consortia of industrially polluted soils

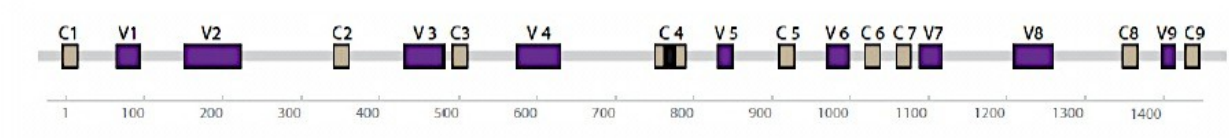
Soil is one of the most species-rich and diverse environments on our planet, including members from all three domains of life: Bacteria, Archaea and Eukarya (Wall et al., 2010). It is estimated that one gram of soil contains around a billion bacterial cells and thousands to millions of bacterial species (Gans et al., 2005). Microbial communities play an important role in nutrient cycling and degradation of contaminants. An improved understanding of the diversity that exists below-ground can help explain the ecological mechanisms underlying community structure and develop conceptual models of the factors controlling microbial diversity, as microbial ecology still lags behind plant and animal ecology in our ability to understand and predict distribution patterns.

For many years the methods of exploring soil biodiversity relied solely on cultivation of bacteria on artificial media (Conn et al., 1918). Later, cultivation-independent methods confirmed the suspicion that the cells that were forming colonies were unrepresentative of the total bacterial community. It has been estimated that 0.1 – 1% of the total soil population can be cultured applying standard cultivation techniques (Kellenberger, 2001). These uncovered microorganisms represent an unexplored reservoir of novel strains, which among other can be a source of novel products (Daniel, 2004).

Metagenomics, a term first suggested by Handelsman, implies habitat-based investigation of mixed microbial populations at the DNA level (Handelsman et al., 1998). It involves the application of molecular ecological methods, especially those based on exploration of genes after PCR amplification and allows for cultivation-independent studies of the microbial communities of soils. One of the main areas of research from the beginning was the discovery of novel biocatalysts.

In time, metagenomic approaches began to play an important role in linking population diversity to environmental processes. Thus, apart from targeting functional genes, metagenomics also studied phylogenetically informative genes and RNAs. In particular, the genes for 16S rRNA have proven to be a useful marker of microbial taxa (Pace, 1997). The 16S rDNA includes interspersed conserved and variable regions (see Figure 1), which makes this gene suitable for PCR amplification and sequencing. In most cases, the probes are designed to hybridize to the conserved regions, allowing for amplification and sequencing of the variable regions.

Figure 1: The schematic location of variable (purple) and conserved (brown) regions in a canonical bacterial 16S rRNA gene. The black region is invariable in all bacteria. Source: Illumina publication reviews, http://www.illumina.com/content/dam/illumina-marketing/documents/products/research_reviews/metagenomics_research_review.pdf



This approach has been particularly effective for monitoring fluctuations in microbial populations (Caporaso et al., 2012). Focusing on a small part of the microbial genome allows for comparison studies and lowers the costs of sequencing. Since the first pilot studies to narrate soil bacterial communities using molecular ecological surveys, a number of libraries of 16S rRNA genes derived from soils have become available: from US grasslands (Zhou et al., 2003) to Swiss pastures (Marilley and Aragno, 1999) and Brazilian forests (Borneman and Triplett, 1997).

Gene exploration through sequencing has evolved from the introduction of dideoxynucleotide triphosphates (ddNTPs) as DNA chain terminators by Nobel Prize laureate Frederick Sanger (Sanger, 1975), through further upgrade with fluorescently labeled ddNTPs introduced by Hood and coworkers (Smith et al., 1986), to the use of high-resolution capillary electrophoresis separation (Kasper et al., 1988).

With the advent of capillary electrophoresis -based Sanger sequencing, researchers gained the ability to elucidate genetic information from any given biological system. This technology has become widely adopted in laboratories around the world, yet has always been hampered by inherent limitations in throughput, scalability, speed, and resolution that often prevented scientists from obtaining the essential information they need for the research. To overcome these barriers, a new technology was implemented — Next-Generation Sequencing (NGS), with its massively parallel approaches bringing definite cost advantages (Mukhopadhyay, 2009).

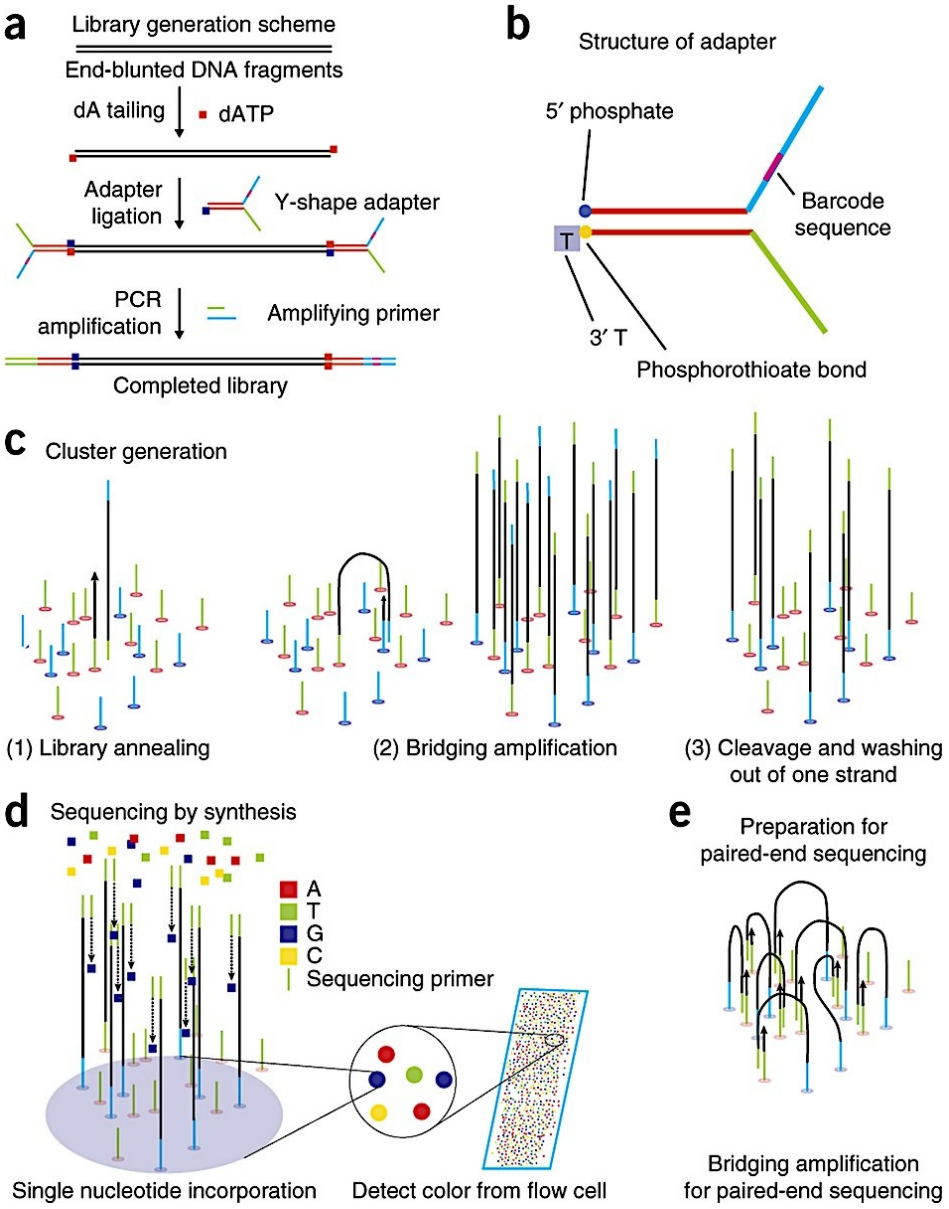
Among the NGS platforms, Illumina is currently the dominant vendor of high-throughput DNA sequencing machines. According to the OmicsMaps website (http://omicsmaps.com/stats, 08.04.2015), to date, Illumina has 6203 machines in place worldwide compared to 1186 for all other vendors (Roche 454, ABI SOLiD and other) combined.

The Illumina sequencing method is similar to Sanger sequencing, but it uses modified dNTPs containing a terminator which blocks further polymerization - so only a single base at a time can be added by a polymerase enzyme to each growing DNA copy strand. The sequencing reaction is conducted simultaneously on a very large number (tens of millions of clusters, according to Illumina documentation) of different template molecules spread out on a solid surface. In contrast to the 454 and ABI methods, which use a bead-based emulsion PCR to generate "polonies" (polymerase generated colonies), Illumina utilizes the so-called "bridged" amplification reaction that occurs on the surface of the flow cell (basically a water-tight microscope slide).

During each sequencing cycle, a single labeled deoxynucleoside triphosphate (dNTP) is added to the nucleic acid chain. The nucleotide label serves as a terminator for polymerization, so after each dNTP incorporation, the fluorescent dye is imaged to identify the base and then enzymatically cleaved to allow incorporation of the next nucleotide. Since all four reversible terminator-bound dNTPs (A, C, T, G) are present as single, separate molecules, natural competition minimizes incorporation bias. Base calls are made directly from signal intensity measurements during each

cycle, which significantly reduces error rates compared to other technologies. The end result is accurate base-by-base sequencing that eliminates sequence-context specific errors, including repetitive sequence regions and within homopolymers.

Figure 2: The scheme of next-generation sequencing process using Illumina platform (Shin et al., 2014)



a. The library is composed of fragments of double-stranded DNA that can be recognized by the NGS sequencer. DNA to be sequenced is flanked by adapter sequences. The DNA to be sequenced should first be fragmented down to 200–600 bp (there is no need for this step in the present study, as an amplicon library of 300-bp fragments is sequenced), and fragment ends should be enzymatically blunted. A 3’ dA tail is then added using DNA polymerase without 3’-5’ exonuclease activity. This dA tailing prevents the formation of concatemers of DNA fragments and allows for the use of a

dT-tailed adapter, which minimizes adapter dimer formation. An Y-shaped adapter is then ligated using A-T base pairing, and the correctly ligated libraries are amplified.

b. The 5' end of the Y-shaped adapter contains a phosphate group and the 3' end contains dT. Phosphorothioate bonds provide resistance to nuclease.
c. Cluster generation on the surface of the flow cell.
    1. Denatured libraries are annealed to the short oligonucleotides on the surface of the flow cell. The distance between DNA molecules should be long enough to prevent overlapping clusters.
    2. Bridging amplification generates clusters.
    3. One strand from the double-strand DNA library is cleaved and washed out for unidirectional sequencing.
d. Primers for inserts are annealed for the sequencing of the insert DNA. In each sequencing cycle, protected and fluorescently labeled A, T, G and C bases are applied. After the addition of each nucleotide, the sequencing reaction is stopped, and the image is taken. Because the newly added nucleotides within each cluster are identical, the signal is high enough to be detected by a light sensor. After the image is taken the protection group and the fluorescent molecules are removed.
e. When the first-strand sequencing reaction is finished, the synthesized strand is removed and the process is repeated for the opposite strand.

Each new read contains both the fragment sequence and its sample identifying barcode. Barcode sequences are subsequently used to de-multiplex or differentiate reads from each sample. The assembly of paired-end reads (see Figure 10 in Methods section) offers an advantage by incorporating an additional quality control step for each 16S rDNA sequence.

Molecular identification of microorganisms and population dynamics studies have been conducted in various ecological niches, for instance: in marine salinity meromictic lakes (Bowman et al., 2000), in reactor treating organic household waste (Cardinali-Rezende et al., 2009), freshwater sediments in response seawater intrusion (Edmonds et al., 2009), in geothermal springs (Kormas et al., 2009) and glacier ice (Simon et al., 2009), most recently in temperate forest soils (DeAngelis et al., 2015) and iron mining area soils (Hong et al., 2015).

The vast majority of soil diversity studies have been conducted in natural (or eventually clinical) settings, yet anthropogenic pressures now influence many ecosystems and it remains unclear whether there are consistent processes structuring soil biodiversity and biogeography in natural and urban ecosystems. As you will see below, the release of antibiotics, as such, into the natural environment has also been extensively studied in the last years. However, to date, the metagenome microbial community dwelling in pharmaceutically polluted soils has never been explored.

Release of antibiotics into the environment is considered to lead to adverse ecotoxicological effects, as well as to the maintenance and selection of antibiotic-resistant bacteria, even in the environments with very low antibiotic concentrations (Gullberg et al., 2011; de la Torre et al., 2012). Little is known about the biodegradability of antibiotics in natural environment and their role in growing bacterial resistance. Biodegradability of cefotiam, ciprofloxacin, meropenem, penicillin G and sulfam ethoxazole in aquatic environments was investigated using a closed bottle test. None of the test compounds met the criteria for ready biodegradability. Only penicillin G was biodegradable to some degree (27%), even when the test was prolonged from 28 to 40 days (35%) (Al-Ahmad et al., 1999). As concerns the persistence in soil, penicillin belongs to impersistent class of compounds and cephalosporin derivatives to moderately persistent with a half-life of 22 - 49 days (Gilbertson et al., 1990).

Furthermore, mixtures of several antibiotic compounds can have synergetic toxic effect on bacteria. It was concluded that antibiotic drugs emitted into environment affect the biological processes, may persist there and contribute to the increasing resistance of pathogenic bacteria (Al-Ahmad et al., 1999). These results are in consistence with other studies investigating the biodegradability of clinically important antibiotics in soil (Marengo et al., 1997; Weerasinghe and Towner, 1997).

A recent study conducted by a Spanish team, performed a risk analysis of the EU region for twelve antimicrobials (de la Torre et al., 2012). The antibiotics in scope belonged to seven chemical groups: tetracyclines, sulphonamides, fluoroquinolones, phenicols, beta-lactams (represented by amoxicillin), macrolides and lincosamide. A risk assessment was carried out to measure soil vulnerability to antibiotic contamination. The assessment was comprised of four steps: release assessment (how much of each active substance is used), exposure assessment (see below), consequence assessment (the geo-referenced information about land uses), and risk estimation (integrating release, exposure and consequence layer). Exposure was estimated as the antibiotic potential to contaminate soil. The potential to contaminate soil, $C=(B+P)/2$, was evaluated for each antibiotic as a function of its binding rate (B) and permanence rate (P). The antibiotic of highest risk was enrofloxacin, followed by tetracyclines, tylosin and sulfodiazine. Beta-lactams, with relatively low persistence rate in comparison to tetracyclines or fluoroquinolones, but high adsorption coefficient, presented eminent contamination potential for instance for France, The Netherlands, Belgium and Ireland.

The study concluded, that the mean risk value (in relative units) in the EU was 0.089, with The Netherlands (0.397), Ireland (0.381), Belgium (0.305), Luxembourg, Denmark, Germany and the UK being the countries at highest risk. Bulgaria (0.018), Greece (0.015) and Sweden (0.014) were at lowest risk. Czech Republic, with 0.063±0.010 soil vulnerability risk, took the 16th position

out of 34 countries in scope. Similar investigations concerning antibiotic pollution were conducted in China (Xue et al., 2013), Southeast Asia (Suzuki and Hoa, 2012) and other parts of the world.

Antibiotics are often used to prevent sickness and improve production in animal agriculture, and their residues may enter wastewater treatment plants. A research group from Nanjing University in China used Illumina high-throughput sequencing to investigate the occurrence, diversity and abundance of antibiotic resistance genes in aerobic and anaerobic sludge of a full-scale tannery wastewater treatment plant. They found out that the genes involved in virulence, disease and defense occupied around 3% of the total reads, of which resistance to antibiotics and toxic compounds occupied over 60%. Specifically in the anaerobic sludge, sulfonamide, tetracycline and multidrug resistance genes had the highest abundance, each around 25% of all resistance genes; with the prevalence explained by frequent use of tetracycline and sulfonamides for livestock purposes in China and high persistence of these antibiotics in the environment (Wang et al., 2013).

Nonetheless, it was demonstrated by other research teams, that certain resistance genes found in soil bacteria are only distantly related to those found in clinical isolates, indicating that the relationship between these genes is not yet fully explored (Donato et al., 2010). The role of the environmental reservoirs in clinical resistance development is still hypothetical. There is little evidence that any of the putative resistance genes identified in the environmental studies have been mobilized into pathogenic bacteria and expressed as resistance phenotypes (Schmieder and Edwards, 2011). Also, there are studies where no shifts in microbial community structure and increased antibiotic resistance were observed in response to antibiotic contamination, suggesting that the soil microbial communities were robust to the effects of antibiotics at test concentrations, e.g. (Unger et al., 2013). However, in the abovementioned study no metagenomic methods were applied - the results were acquired by Phospholipid Fatty Acid analysis and broth microdilution.

It should be mentioned, that there are certain methodological limitations of the present study. First of all, the research focuses solely on bacteria, with no attention paid to unicellular and multicellular eukaryotes. However, there is increasing evidence that the diversity of soil fungi, protists and metazoa is likely far higher than often considered, and their influence on prokaryotic community is quite significant. This influence can be a result of trophic interactions between predatory protists and their bacterial prey, a product of direct symbioses (e.g. the relationships between fungi and bacteria), or simply shared environmental drivers. Only a few cross-domain assessments of soil diversity exist to date, e.g.: (Fierer et al., 2007; Delmont et al., 2012).

Other unmeasured variables (humidity, temperature, salinity etc.) may also contribute to the observed biological patterns. Soil pH has been shown to be an important driver of bacterial community composition (Lauber et al., 2009; Rousk et al., 2010). The most probable mechanism of the pH impact on microbial consortia may include mediation of nutrient availability in the soil. A

different theory implies a narrow pH tolerance of some prokaryotic taxa. Acidification of the interior of a microbial cell inhibits the activity of most enzymes and of overall cell metabolism, however, many soil microorganisms have developed adaptive responses to survive in acidic environments. In any case, the highest bacterial diversity was found in neutral soils with pH around seven (Zhalnina et al., 2015).

These conclusions are in a good agreement with the results of the study based on the Park Grass Experiment. It was established in 1856 and is claimed to be the oldest ecological experiment in the world. The experiment has allowed scientists to study the influence of nutrient additions and different pH levels on biodiversity and ecology for many years. It revealed, among other, a decrease in bacterial and an increase in fungal growth with low pH levels and a strong positive correlation between the abundance of bacterial operational taxonomic units (OTUs) and soil pH (Rousk et al., 2010; Rousk et al., 2011).

Indeed, some research groups came to a conclusion that no other environmental variable (e.g. (humidity, fertilizers, including nitrogen as ammonium sulphate, nitrogen as sodium nitrate, phosphorous) is significantly correlated with the prokaryotic community distribution patterns. Interestingly, soil pH was the only measured variable that significantly correlated with eukaryotic biogeographic patterns, but it was not a particularly strong predictor. Instead, the best predictor of eukaryotic community patterns was prokaryotic community composition, which again emphasizes the importance of cross-domain studies. (Ramirez et al., 2014). Another study suggests that there is no clear relationships between plant diversity and soil bacterial diversity at the continental scale (Fierer and Jackson, 2006).

It is important to decide at which level to assess the metagenomic biodiversity. It was found that samples coming from ecologically similar environments with similar phylum-level 16S rRNA gene distributions could contain almost completely distinct collections of biosynthetic gene sequences. Most probably, this is due to the fact that natural product biosynthesis gene content can differ not only between species but also between strains of the same species (Reddy et al., 2012). This can be one of the reasons why most researchers are focused either on biodiversity (16S rDNA) or metabolic analysis. Some studies suggest that this connection can only be made by reconstructing partial or entire organisms from the community through metagenomic assembly (Liu et al., 2012). As a result of cross-biome metagenomic analysis of soil from deserts, forests, grassland and tundra, it can be stated that at least low functional diversity was found to correlate with low taxonomic diversity (Fierer et al., 2012).

Additionally, the so-called mosaicism can lead to misidentification. Many bacteria have a history of horizontal gene transfer, and can tolerate transfer of complete 16S genes. And vice versa: horizontal transfer of functional genes, or even significant genomic rearrangements, may not be reported by the 16S rRNA region (Asai et al., 1999; Altermann, 2012).

Many comparison studies do not take into account, that bacterial genome can contain more than one copy of rRNA operon. Actually, the number of copies varies from 1 to 15 per bacterial genome (possibly reflecting different ecological strategies of bacteria (Klappenbach et al., 2000)), and the sequences of multi-copy rRNA can vary by 6.4% of total nucleotides (Wang et al., 1997). This can impact abundance estimates based on the rRNA and limit the phylogenetic resolution of species based on those sequences. In the present study, to reduce this potential impact, read abundances of each OTU were divided by the copy number of the rRNA genes in the genome of the closest taxon with a complete genome sequence (Větrovský and Baldrian, 2013).

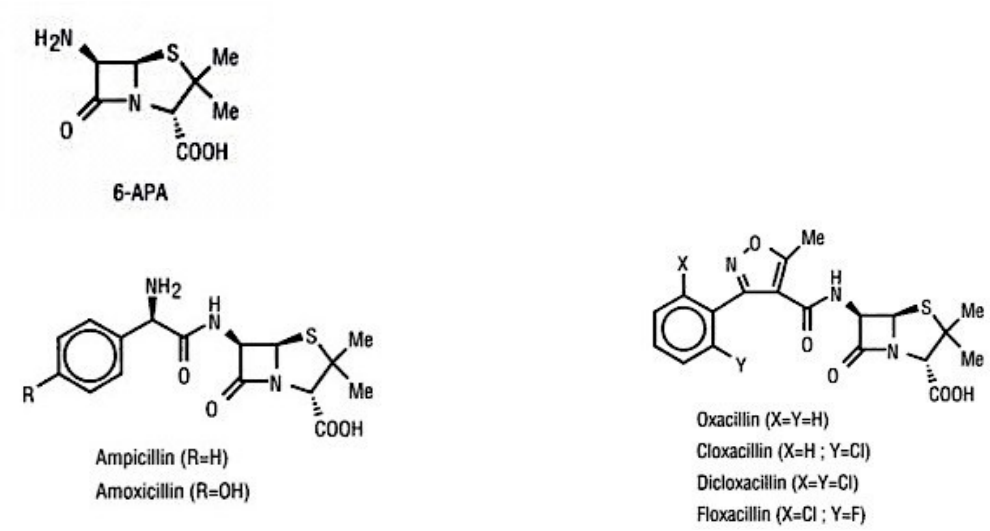
Illumina next-generation sequencing is a very accurate method, although it is inevitably influenced by PCR and cloning biases. For instance, the primers for 16S rDNA fragment amplification are designed to hybridize to conserved regions, but these regions may change in evolution, resulting in a loss of hybridization to the probe, which in its turn leads to an underestimation of evolutionarily distant members of the population. Furthermore, the variable regions are of different sizes and can change at different rates; it has been concluded that V6 tags (see Figure 1) appear to systematically overestimate species richness (Youssef et al., 2009). Paired-end Illumina sequencing of the larger V4 region, applied in the present study, has been proven to successfully build phylogenetic trees (Werner et al., 2012).

Nevertheless, in general, metagenomic approach has been quite successful in helping to provide a more comprehensive description of the organisms and processes that shape microbial community structure, adaptability and dynamics in different ecosystems, e.g. (DeLong, 2005). It should be mentioned though, that in the past, in the absence of metagenomic techniques, pure cultures derived from cultivable soil microorganisms provided much of our basic knowledge of soil bacteria, allowed for the discovery of many important antibiotics and useful enzymes (Alexander, 1978). Cultured isolates are still very important in developing our understanding of bacterial physiology, genetics, and ecology (Janssen, 2006). Therefore the second part of the presented study also deals with cultivable organisms isolated from soil, screening among these pure cultures and subsequent gene-mining.

1.2 Syntheses of beta-lactam antibiotics: utilization of AEH and other enzymes

Most natural products of industrial potential and economic value including antibiotics and other pharmaceuticals are derived from cultured soil microorganisms (Daniel, 2004). The antibacterial activity of *Penicillium notatum* was first discovered by Alexander Fleming in 1928, when the scientist realised that this fungus is able to inhibit the growth of *Staphylococcus aureus*. (Fleming, 1929). It took 12 years before Florey and Chain eventually isolated penicillin G due to the relative instability of its nucleus (Florey, 1949). Their success was followed by description and wide medical application of this antibiotic, until it was phased out of clinical use in the late 1960's after many strains became resistant to this type of penicillin (Wegman et al., 2001). The problem was resolved with the help of semi-synthetic antibiotics, in which the beta-lactam nucleus of penicillin G – 6-aminopenicillanic acid (6-APA) – is modified with various side-chains as illustrated in Figure 3.

Figure 3: The structure of 6-APA (6-aminopenicillanic acid) as penicillin nucleus, and major penicillins (Bruggink, 2001)

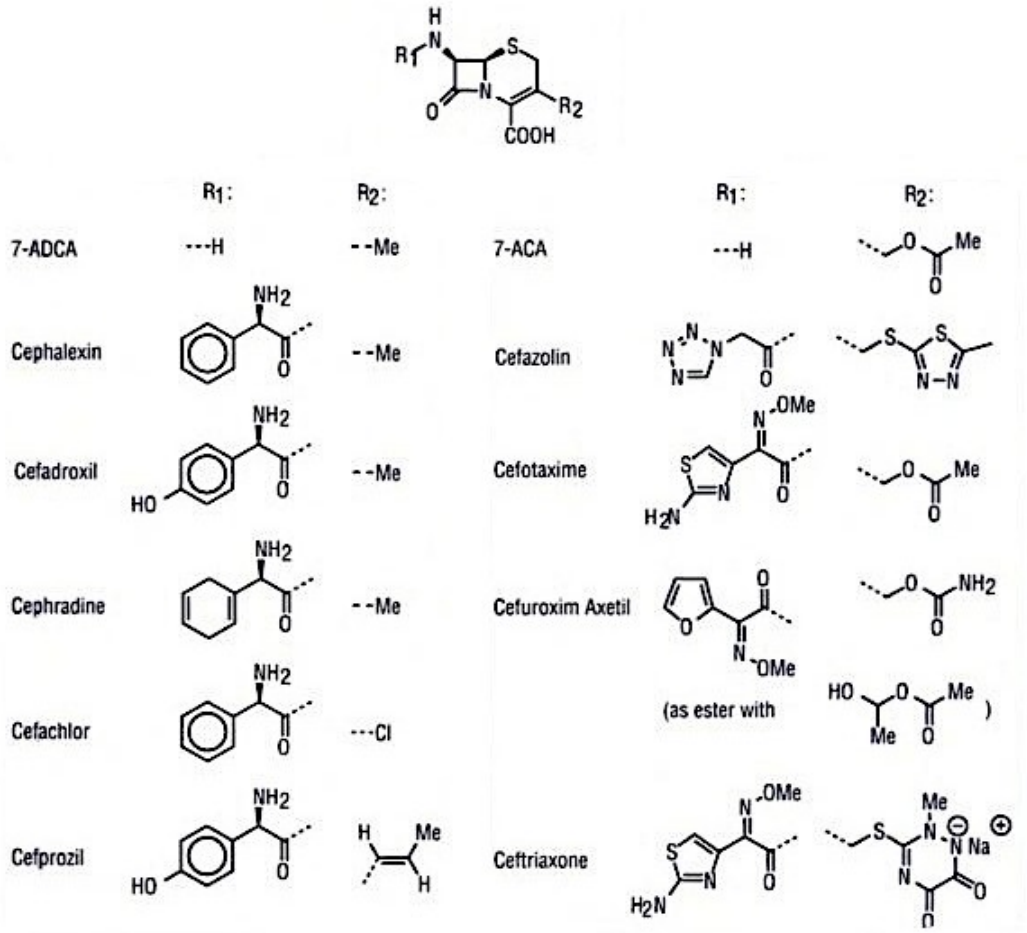


A new type of beta-lactam antibiotics was discovered in 1948 (Brotzu, 1948), when Giuseppe Brotzu isolated cephalosporin C from fermentation broth of *Acremonium chrysogenum* (previously known as *Cephalosporium acremonium*). The fungus showed high antagonistic potency against *Staphylococcus aureus*, *Vibrio cholera*, and *Bacillus anthracis*. Structurally, the main difference between penicillins and cephalosporins is that the nucleus of the former consists of a five-membered heterocycle fused to a four-membered, whereas the skeleton of the latter contains a six-membered ring. (Newton and Abraham, 1955).

Due to the relatively low efficacy of naturally occurring cephalosporins produced by *A. chrysogenum*, all clinically important cephalosporins are manufactured semi-synthetically either from 7-ACA (7- aminocephalosporanic acid) or from 7-ADCA (7-aminodeacetoxycephalosporanic

acid) as a nucleus (Harris et al., 2009). Cephalosporins are relatively resistant to beta-lactamases and exhibit potency against both gram-positive and gram-negative bacteria (Weil et al., 1995). Additionally, as compared to penicillins, which possess a single site for derivatization, cephalosporins offer two sites: the 7-amino group and the C3 acetoxy group (see Figure 4) allowing for production of a larger number of semi-synthetic derivatives.

Figure 4: The structure of cephalosporin nuclei 7-ACA (7- aminocephalosporanic acid) and 7-ADCA (7-aminodeacetoxycephalosporanic acid), and important cephalosporins (Bruggink, 2001)



Some time later a new family of beta-lactam antibiotics was discovered, known as cephamycins. These antibiotics are produced only by prokaryotes of the actinomyces genus, are characterised as cephalosporins modified at C7 or C3 (or both) and are especially active against gram-negative bacteria (Demain and Elander, 1999).

It should be mentioned that apart from penicillins (penams) and cephalosporins/cephamycins (cephems), that typically contain sulphur in heterocycle, there exist other groups of beta-lactam antibiotics: clavams, monolactams and carbapenams (Brakhage, 1998). Mechanism of action of almost all of the above-mentioned antibiotics can be briefly described as inhibiting bacterial cell wall biosynthesis by interacting with penicillin binding proteins, which synthesize and remodel peptidoglycan, the structural component of bacterial cell walls (Denome et

al., 1999). This results in lysis of the growing bacterial cells, which do not bear the genes encoding beta-lactamase or possess any other mechanisms of drug resistance. It should be noted, that emerging resistance of bacteria to antibacterial drugs, among other under pressure of natural selection during antibacterial drug therapy, is a major challenge nowadays (Hayes and Wolf, 1990).

According to a recent market report from Transparency Market Research (http://www.transparencymarketresearch.com/antibiotic-market.html 30.04.2015), a global market intelligence company, the global antibacterial drugs market was valued at $43.55 billion in 2012 and is expected to reach an estimated value of $45.09 billion in 2019. The beta-lactams, being the largest segment of the global antibacterial drugs market in 2012, have a 57% share. Cephalosporins are one of the most widely used antibiotics and comprise approximately 29% of the antibiotic market. Due to constantly growing production efficiencies, bulk production costs for the above mentioned beta-lactams were moved downscale, in contrast with many other marketed pharmaceuticals. At the same time the labour and energy costs increased significantly. These factors contribute to the trend towards relocation of antibiotic manufacturing to the Far East and other developing countries, with India becoming the largest generic drug-manufacturing nation. Nevertheless, continuing research especially concerning new generation cephalosporins is still justified in Europe, United States and Japan (Barber et al., 2004). There is a need for further improvement of production routes towards existing antibiotics, as well as for more effective antibiotics against resistant pathogens. The variety of combinations between beta-lactam nuclei and synthetic side-chains allows for the development of such pharmaceuticals.

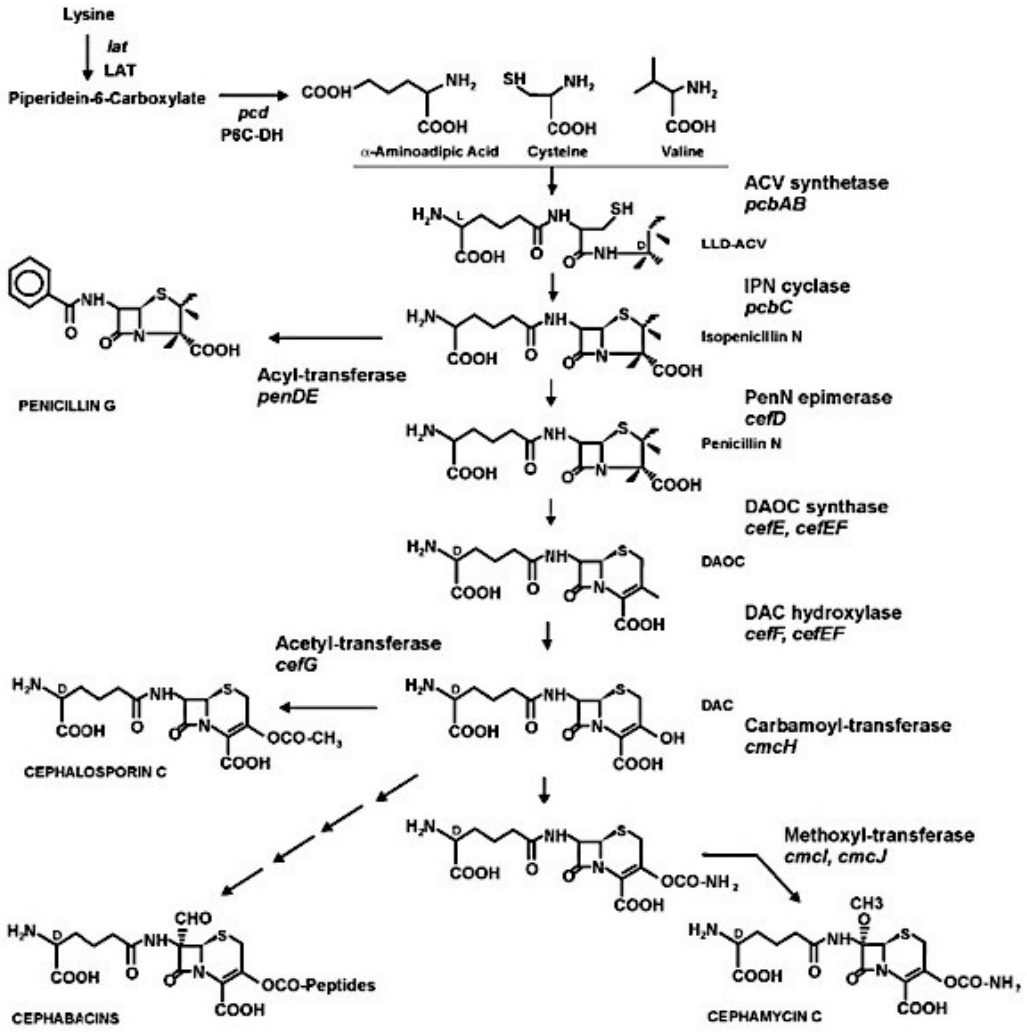
1.2.1 Syntheses of natural beta-lactams

Naturally occurring biosynthetic pathways towards penicillins and cephalosporins are well characterised genetically and biochemically (Demain, 1983; Baldwin and Abraham, 1988; Martin and Liras, 1989; Martin and Gutierrez, 1995). The pathways start identically, as can be seen in generalized scheme in Figure 5. The first step is the production of L-delta-(alpha-aminoadipoyl)-L-cysteinyl-D-valine (ACV) by ACV synthase. This step is followed by conversion of ACV into isopenicillin N by isopenicillin N synthase (IPNS). At this point the routs towards different beta-lactam antibiotics diverge. In fungus *Penicillium chrysogenum*, which is a major industrially used producing organism of penicillin G, IPN is then converted by isopenicillin N acyltransferase into penicillin G. Alternatively, in cephalosporin producing organisms, such as *Acremonium chrysogenum,* and bacterium *Streptomyces clavuligerus* IPN is converted to penicillin N by an isomerase. Further, the five-membered thiazolidine ring of penicillins is expanded to six-membered dihydrothiazine ring of cephalosporins by an enzyme called deacetoxycephalosporin C synthase (expandase or DAOC synthase). Continuing in the pathway, deacetoxycephalosporin is

hydroxylated and acetylated to produce cephalosporin C. Additional reactions are involved during cephamycin synthesis pathway in the actinomycetes: the attachment of methoxyl and carbamoyl group (Martin and Liras, 1989).

The genes encoding the key enzymes involved in biosynthesis of penicillin G in *P. chrysogenum* are in a cluster (Thykaer and Nielsen, 2003). The same is true for cephamycin synthesis in *S. Clavuligerus* (Ward and Hodgson, 1993). In *A. chrysogenum* two genes involved in natural biosynthetic routes towards cephalosporins - *pcbC* encoding isopenicillin N-synthetase and *cefEF* encoding deacetoxycephalosporin C synthase/hydroxylase - are located on a different chromosome than the rest of the genes involved (Skatrud and Queener, 1989).

Figure 5: Natural biosynthetic pathways towards penicillins and cephalosporins (Liras and Martín, 2006)



1.2.2 Beta-lactams: biocatalytic transformations as compared to chemical synthesis

As was already mentioned, naturally produced penicillin G could not be used clinically anymore due to the widely spread resistance to it, and cephalosporin C almost lacks antimicrobial activity. To obtain medically exploitable beta-lactams either biocatalysis or chemical synthesis can be applied. The biocatalysis here is understood as a chemical reaction catalysed by an enzyme, leading to a conversion of one functional group or structure in a substrate into a different group or pattern in a product (Martin et al., 2009).

1.2.2.1 Chemical synthesis

In the past predominantly chemical syntheses used to be applied.

Thus, to produce semi-synthetic penicillins, penicillin G (available from fermentation process) needed to be deacylated in 4 chemical steps to acquire its nucleus 6-APA. Briefly, the selective cleavage of the secondary amide bond was achieved by its transformation into imine chloride, using phosphorous pentachloride in dichloromethane, followed by hydrolysis. A very important step was the protection of the penicillin carboxyl group by a silyl group. The process was proposed by Weissenburger and van der Hoven in the year 1970, remained in use for 15-20 years and is known as "Delft Cleavage". This route towards 6-APA was inexpensive, but dependent on hazardous, highly active chemicals, that were additionally difficult to recover or reuse, generating excessive amount of waste (Wegman et al., 2001).

Afterwards, semi-synthetic penicillins, e.g. ampicillin - the first to be commercialized - were obtained by either acid chloride route (included silylation and low temperatures) or the Dane salt anhydride procedure (Dane and Dockner, 1964). The latter provided the product in excellent yield (Henniger, 1978) and became widely adopted until the need for a more efficient process was recognized in the early 1990's. The reason was excessive waste, a large part of which was not biodegradable and originated in the activation, coupling and deprotection steps (Wegman et al., 2001).

As concerns the chemical production of cephalosporins, about one-third of these antibiotics is derived from 7-ADCA and two-thirds from 7-ACA as a nucleus (Elander, 2003).

Acquiring 7-ADCA by the removal of side-chain from deacetoxycephalosporin C is not trivial, because no suitable enzyme was found to perform this cleavage. However, in the year 1963 a chemical correlation between cephalosporins and penicillins was explained by Morin et al (1963). The penicillin nucleus could be chemically expanded with the use of sulfoxidation of penicillin G, esterification and subsequent dehydration leading to formation of 7-ADCA from 6-APA (Morin et al., 1963). The process for 7-ADCA production, as it is generally applied today, also involves oxidation of penicillin G and esterification. The dried product is then silylated,

and in the presence of hydrogen bromide catalyst and pyridine a ring expansion is achieved. The protective silyl group is subsequently removed by water addition providing cephalosporin G formation. This intermediate can be afterwards enzymatically cleaved by PGA resulting in 7-ADCA formation (Bruggink, 2001).

The other important type of cephalosporin nucleus, 7–ACA, is produced by the removal of the 7-aminoadipoyl side-chain from cephalosporin C. The chemical, nitrosyl chloride, cleavage was elaborated by Morin (Morin et al., 1969), but has later been superseded by a modified procedure, in which after the amino and carboxyl reactive groups are protected, cephalosporin C is cleaved with the use of phosphoric pentachloride (Barber et al., 2004).

A semi-synthetic antibiotic derived from 7-ACA as a precursor, for instance cephadroxyl, can afterwards be chemically synthesised using silylation followed by an acylation with a mixed anhydride prepared from a salt of p-D-hydroxyphenylglycine and ethylchloroformate (Bruggink, 2001). Alternatively, both cleavage and acylation processes can be performed enzymatically with certain benefits as well as disadvantages, as will be discussed below.

1.2.2.2 Biocatalytic synthesis

Today, with the volumes of beta-lactam antibiotics produced and the general trend towards ecologically friendly technologies, the benefits of biocatalytic processes became obvious.

1. Safety and environment preservation: biocatalytic procedures take place in aqueous environments, dangerous or toxic reagents are absent in the enzymatic transformation. Thus, according to DSM (https://www.dsm.com/content/dam/dsm/cworld/en_US/documents/factbook-2012.pdf), the implementation of biocatalytic procedures in beta-lactam production reduced process greenhouse emissions by over 50%, emissions to water by 50% and process waist by up to 90%.
2. Selectivity: owing to enzyme selectivity, there is also no need for extensive protection of functional groups of reagents and their activation.
3. Energy consumption: the chemical process uses low temperatures and exothermic steps, while the enzymatic approach is carried out at neutral pH and around room temperatures.
4. Equipment: the implementation of biocatalysis was supported by the fact that it was possible to run enzymatic processes in existing equipment, with enzymes being part of organic chemist's toolbox (Bruggink, 2001).
5. Quality: apart from ecological and economical benefits biocatalysis provides high stereoselectivity (Lee et al., 2009).

The original drawbacks of the use of enzymes in biocatalysis were their costs, difficulties in handling and applicability solely in aqueous solutions. At present these disadvantages have been overcome mainly owing to extensive screening, recombinant DNA technology (Wegman et al., 2001) and enzyme immobilization (Hegde et al., 1997).

As concerns the application of biosynthetic methods in industrial practice, as early as in the year 1960 four pharmaceutical companies (Bayer, Beecham, Bristol and Pfizer) independently announced the use of penicillin G acylases (PGA). Once the robust, immobilized biocatalysts reached the market (which made their recycling possible), biosynthetic routes began to dominate over chemical ones in pharmaceutical industry.

Thus, immobilized penicillin acylase as a biocatalyst is used for enzymatic synthesis of cephalosporin type of antibiotics, e.g. coupling of 7-ADCA and D-phenylglycine amide (as a side-chain) to produce cephalexin (see Figure 4). The process does not require protective groups, activating agents, extreme temperatures and can be carried out in water. The procedure however suffers from several drawbacks, such as low coupling efficiency due to secondary hydrolysis of the product and difficult down-stream processing from a complex mixture of chemically similar molecules - amino acid derivatives (Bruggink, 2001). Although the latter has been solved by adding beta-naphthol to the solution and its precipitation with cephalexin (Faarup, 1977), the economic profit is still difficult to obtain as compared to chemical coupling.

The manufacturing of 7-ACA *via* enzymatic cleavage of alpha-aminoadipyl chain from cephalosporin C in contrast with deacylation of penicillin G is not efficient. The activity of known suitable enzymes are low and no appropriate technology for a one-step biocatalytic conversion has been developed yet (Parmar et al., 1998). However, a two-step cleavage with D-amino acid oxidase (DAAO) and glutaryl acylase (GAC) (Arnold, 1968) is industrially acceptable and widely used (Barber et al., 2004).

The conversion of 7-ACA into an active antibiotic by adding a new side-chain can be performed in an enzymatic process catalysed by penicillin G acylase (Kasche et al., 1987). Alternatively, both the removal of the 7- amino adipoyl side chain and the addiction of a new one can be achieved through chemical route.

Today the use of penicillin G acylases for enzymatic cleavage of penicillin G has universally replaced the “Delft Cleavage” process (Wegman et al., 2001). The same enzyme is capable of subsequent acylation of the 6-APA (or 7-ADCA for cephalosporins) nucleus with an acyl donor to produce SSBA. (Bruggink et al., 1998). The biocatalysis of beta-lactams is a very efficient process nowadays, owing to classic improvement techniques, such as random mutagenesis, temperature, pH and oxygenation control, but additional refinement is still justified.

1.2.2.3 Current state of the art of beta-lactam production by cell factories

Improvement of penicillin G fermentative production by recombinant *P. chrysogenum*

To improve the penicillin production, additional copies of *pcbC* (encoding IPN synthase) and *penDE* (encoding acyl transferase) genes were introduced (Veenstra et al., 1991) into *P. chrysogenum*. Later, another study was performed on a single copy strain of *P. chrysogenum* showing that the largest increase in specific penicillin productivity was in transformants with the whole penicillin gene cluster amplified (Theilgaard et al., 2001). Further research revealed that highly productive industrial strains of *P. chrysogenum* previously improved by random mutagenesis contain several penicillin gene clusters.

However, there seems to be an upper limit of five clusters, beyond which the relationship between the copy number and productivity is not linear. This is most likely due to limitations in primary metabolism and a shift to a more complex flux control in a cell (Thykaer and Nielsen, 2003).

Improvement of cephalosporins enzymatic production by recombinant *P. chrysogenum*

At present major effort is put into improving cephalosporin production processes as a more lucrative field of study. To give an example, the production costs of 7-ACA are estimated to be four to five times higher than for 6-APA (which are approximately 15-25 $/kg). The reasons are lower yields, expensive purification (including column chromatography) and less efficient enzymatic conversion (Barber et al., 2004).

As was stated earlier, 7-ADCA, as a precursor for semi-synthetic cephalosporins, is currently produced by the chemical oxidative ring expansion of penicillin G to deacetoxycephalosporin G, which is subsequently enzymatically deacylated (Bruggink, 2001). The initial problem of enzymatic production of 7-ADCA is that, by contrast with obtaining 6-APA from *P. chrysogenum,* the removal of the natural D-α-aminoadipyl side-chain from cephalosporin C is inefficient and there is a lack of strains producing sufficient amounts of deacetoxycephalosporin C (Crawford et al., 1995). Moreover, no suitable enzyme was found to perform the cleavage of deacetoxycephalosporin C, which would lead to 7-ADCA production.

As shown in Figure 5, the enzymatic ring expansion occurs naturally in biosynthetic routes of cephalosporin producing organisms. The enzyme which performs the expansion (expandase or deacetoxycephalosporin C synthase) was reported to accept only penicillin N as a substrate, converting it to deacetoxycephalosporin C (Kohsaka and Demain, 1976). Penicillin N, however, is not commercially available. Further studies confirmed that neither penicillin G nor adipyl-6-APA can be expanded by this enzyme (Kupka et al., 1983). Nevertheless, it has been claimed recently, that by proper adjustment of reaction conditions (including concentrations of alfa-ketoglutarate, Fe^{2+}, ascorbate and reducing agents), expandase from *S. clavuligerus* can accept also other

exogenously-added penicillins than penicillin N. In particular penicillin G and semi-synthetic penicillins, e.g. ampicillin, can be exploited (Demain et al., 2002). The invention is subject to a patent US 6,383,773 B2 and could possibly replace the multi-step chemical process of ring expansion, but no reports of its industrial application are known to this date.

The problem of narrow specificity of expandase can be resolved by cloning the expandase (from *S. clavuligerus*) and IPN epimerase genes (from *S. lipmanii*) into *P. chrysogenum*, which normally doesn't produce cephalosporins but is an ultimate penicillin producer, allowing the recombinant strain to produce deacetoxycephalosporin C (Cantwell et al., 1992).

It was also discovered that growing *P. chrysogenum* strain expressing expandase in media containing adipic acid as the side chain precursor would lead to a formation of cephalosporins with adipoyl side-chains. This proved that adipyl-6-APA is a substrate for *S. clavuligerus* expandase enzyme in vivo. Unlike the side-chain in deacetoxycephalosporin C, adipoyl side-chains can afterwards be relatively easily enzymatically removed (several known amidases are capable of such hydrolysis), leading to a production of 7-ADCA (Crawford et al., 1995). It was further disclosed, that not only adipate is a suitable feedstock for an expandase-expressing *P. chrysogenum* strain. Dicarboxylic acids with a chain length, which is longer than 7 carbon atoms present in a cultivation medium lead to a production of acyl-7-cephalosporin derivatives (Nieboer et al., 2003).

In 2009 the route towards adipoyl-7-ADCA was further extended by introduction of *S. clavuligerus* carbamoyltransferase into expandase expressing *P. chrysogenum*. The obtained metabolite, adipoyl-7-aminocarbamoyl cephalosporanic acid represents a new precursor towards several semi-synthetic cephalosporins, including Cefazolin and Ceftriaxone (Harris et al., 2009).

A more recent study was performed in 2007 similar to the experiment carried out by Cantwell et al (1992). The basic assumption was again the suitability of *P. chrysogenum* as a host organism for the production of beta-lactams other than penicillins naturally produced by this fungus. The results of the study indicate that the expression of fungal genes (from *A. chrysogenum*) for the expandase and IPN epimerase instead of bacterial (as suggested by Cantwell et al (1992)) is more rational and provides higher productivity (Ullan et al., 2007).

An inverse approach to what was described by Cantwell et al (1992) is to disrupt the part of the pathway towards unwanted cephalosporin products in *A. chrysogenum*. Thus, instead of introducing desired part of the rout into a new host, a natural cephalosporin producer is genetically improved by disrupting the gene for deacetylcephalosporin C synthase (see the scheme in Figure 5). Together with overexpression of expandase this improvements lead to an increased production of deacetoxycephalosporin C without its contamination with cephalosporin C, which represents an unwanted product in this case (Velasco et al., 2000).

1.2.3 Principle enzymes utilized for biotransformation of beta-lactams

The most widely industrially applied enzymes utilized in the production of semi-synthetic beta-lactams include previously depicted penicillin G acylase (EC 3.5.1.11); D-amino acid oxidase, DAAO (EC 1.4.3.3); and glutaryl acylase, GAC (EC 3.5.1.93). Alpha-amino acid ester hydrolase (EC 3.1.1.43) has been recently described to catalyse syntheses of SSBA. As this thesis focuses on the potential use of AEH as a biocatalyst, all other enzymes are introduced very briefly.

1.2.3.1 Penicillin G acylase

Penicillin G acylase belongs to the family of penicillin acylases, enzymes that catalyse the hydrolysis of penicillins to produce 6-APA and the hydrolysis of 3-deacetoxycephalosporin G to form 7-ADCA. These enzymes also catalyse the reverse reaction: acylation of the amino group of the beta-lactam antibiotic intermediates with appropriate esters or amides leading to the synthesis of semi-synthetic beta-lactam antibiotics. However, penicillin acylases are additionally used for other procedures, such as peptide synthesis, removal of protecting groups and separation of racemic mixtures (Zhang et al., 2006). Apart from PGAs, two other groups of enzymes belong to the penicillin acylase family: ampicillin acylase and penicillin V acylase (Shewale and Sudhakaran, 1997) - the nomenclature is derived from the industrially relevant properties. PGAs, that preferentially hydrolyse penicillin G, have been found in many bacteria, e.g. *Escherichia coli*, *Alcaligenes faecalis*, *Bacillus megaterium*, *Kluyvera citrophila* and *Providencia rettgeri* (Zhang et al., 2006). However the industrial production of penicillin G acylase is conducted in recombinant *E. coli* systems, with a variety of strategies applied to improve their productivity (Yang et al., 2006).

One of the first industrially utilized beta-lactam acylases is the PGA from *E. coli* is a periplasmic enzyme that is composed of an alpha- (small, 209 aminoacids) and beta- (large, 556 amino acids) subunits. The crystal structure revealed a single-amino-acid catalytic centre: Serβ1, located at the terminal end of the beta-subunit (Duggleby et al., 1995). Other well-studied PGAs are also composed of two non-identical subunits, have similar molecular weights and substrate ranges, but vary in kinetic parameters.

1.2.3.2 D-amino acid oxidase and glutaryl acylase

Two main enzymes are important for the production of precursors for semi-synthetic cephalosporins: D-amino acid oxidase and glutaryl acylase. GAC has been isolated from several organisms, but only the enzyme from *Pseudomonas diminuta* expressed in a recombinant *E. coli* is utilized industrially (Barber et al., 2004). Glutaryl acylases are able to hydrolyse glutaryl-7-ACA and less efficiently cephalosporin C to 7-ACA. After unsuccessful attempts to achieve direct one-step transformation of cephalosporin C into 7-ACA with the use of GAC, a two-enzyme process was developed. The step preceding the hydrolysis catalysed by GAC is performed by D-amino acid oxidase, that catalyses an oxidative deamination of cephalosporin C. Subsequently, after

a spontaneous decarboxylation occurs, glutaryl-7-ACA is used to produce 7-ACA (Tischer et al., 1992). DAAOs are known to be produced by numerous microorganisms; the enzymes from yeasts *Rhodotorula gracilis* and *Trigonopsis variabilis* have been developed into industrially applied biocatalysts (Barber et al., 2004).

1.2.3.3 Alpha-amino acid ester hydrolase

AEHs were first described in 1972 by Takahashi *et al.* as enzymes that catalyse the hydrolysis and synthesis of esters and amides with an alpha-amino group. The researchers examined the substrate specificity of the cephalosporin-synthesizing enzyme and classified it as an alpha-amino acid ester hydrolase, due to the fact that:

1. the enzyme was able to catalyze both transfer and hydrolysis reactions,
2. the enzyme preferred esters over amides,
3. the substrate specificity of the enzyme is limited to an acyl donor with an alpha-amino group (Takahashi et al., 1974).

To date, ten strains have been described to produce alpha-amino acid ester hydrolases, the AEH encoding genes from four of them were successfully cloned and expressed in E. coli: *Acetobacter turbidans* (Polderman-Tijmes et al., 2002), *Xanthomonas citri* (Barends et al., 2003), *Xanthomonas campestris pv. campestris* (Blum and Bommarius, 2010) and *Xantomonas rubrilineans* (Pan et al., 2013; Zarubina et al., 2013).

AEH is a homotetrameric periplasmic enzyme with the active site located in each subunit. The subunit is composed of 3 domains: the first domain, domain I (1–197, 285–410) contains the catalytic serine and an α/β hydrolase fold. The cap domain II (198–284) contains mostly helices, while domain III (411–637) makes up the C-terminus (Blum et al., 2012).

According to experimentally determined crystal structures, the presence of three types of key amino acid residues is a characteristic for alpha-amino ester hydrolase (Barends et al., 2003; Barends et al., 2006). Firstly, there can be found a canonical catalytic triad of serine hydrolases, which consists of serine, histidine and asparagine residues. Secondly, there is an oxyanion centre consisting of two tyrosine residues, which is required to stabilize the negative charge on the catalytic serine residue. And finally, there is a carboxylate cluster consisting of three carboxyl groups of two aspartic acid residues and a glutamic acid residue, which is involved in the binding of the positively charged amino-group of the acyl moiety, ensuring high specificity of AEH.

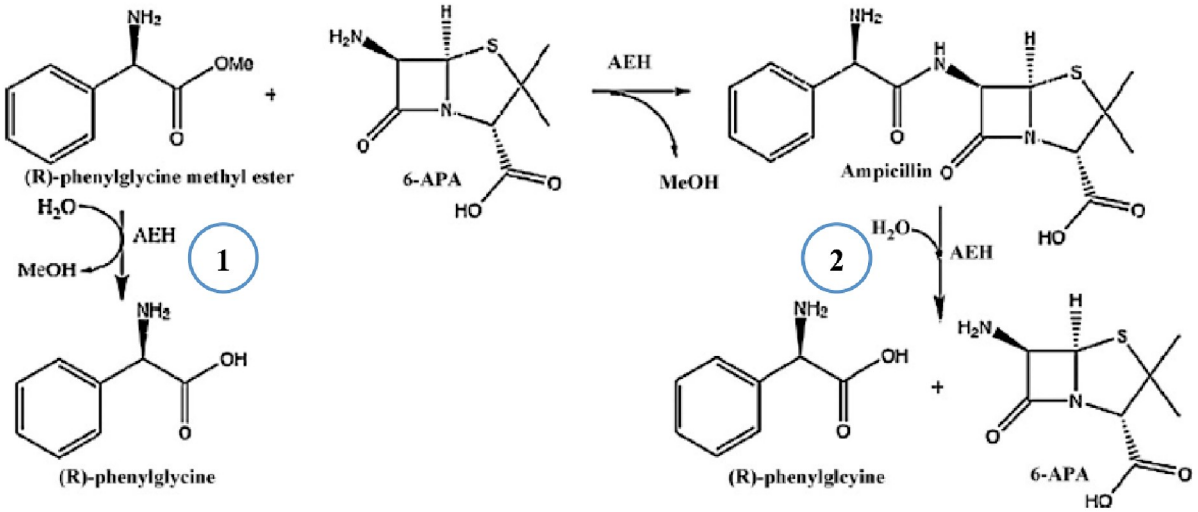
It has been concluded that AEHs catalyse the reactions through a common acyl-enzyme intermediate. In such a mechanism, the substrate first associates with the enzyme to form a noncovalent enzyme-substrate complex (Michaelis-Menten complex) followed by the acylation of the active site serine to give the acyl-enzyme intermediate. This acyl-enzyme is attacked by a nucleophile, to form the enzyme-product complex, from which the product is released. In case of

the AEHs (as well as PGAs) the acyl side chain can be transferred to water or to amine nucleophiles (e.g. beta-lactam nuclei) resulting in hydrolysis or formation of a semi-synthetic beta-lactam antibiotic, respectively (Kato et al., 1980; Blinkovsky and Markaryan, 1993).

It should be mentioned that beside the target reaction AEHs also hydrolyse parent alpha-amino acid esters and the acyl amide bond in the synthesized product. The ratio between the competing synthase, esterase, and amidase activities of the enzyme during the synthesis and hydrolysis of the substrates involved in the reaction scheme of a particular antibiotic is approximately 2 : 5 : 1 respectively (Takahashi et al., 1974; Kato et al., 1980; Kurochkina et al., 2013). However, the AEH specificity to target product hydrolysis is several times lower than to alpha-amino acid ester hydrolysis, which is essential for beta-lactam synthesis.

The biocatalytic synthesis of beta-lactams can be accomplished either *via* thermodynamically (proceeding towards reaction equilibrium) or kinetically controlled reaction. In the case of kinetically controlled synthesis, the yield of condensation product, influenced by the properties of the enzyme, is higher (can compete favourably with the chemical processes) and is also obtained more rapidly (Kasche, 1986). During this procedure, the beta-lactam nucleus is acylated with an activated side-chain donor, as it is illustrated for ampicillin in Figure 6.

Figure 6: Enzymatic synthesis of ampicillin by AEH, adapted from (Blum and Bommarius, 2010). The main drawback of the kinetically controlled biosynthesis of beta-lactams is the secondary hydrolysis of product (2) and hydrolysis of the activated side-chain donor (1) (Boccu et al., 1991)



1.2.4 Comparison of synthetic potential of AEH and PGA

When compared to penicillin G acylase from *E. coli*, alpha-amino acid ester hydrolase has the following advantages:

1. An ability to catalyze beta-lactam synthesis with minimum secondary hydrolysis. In an experiment, carried out by Blum and Bommarius, ampicillin was synthesized through enzymatic acylation of phenylglycine methyl ester (PGME) with 6-APA (as shown in Figure 6) in a purely aqueous system. This reaction is coupled with two enzyme-catalyzed side reactions, primary hydrolysis of PGME to phenylglycine and methanol and secondary hydrolysis of ampicillin to phenylglycine and 6-APA. Similar results for secondary hydrolysis were obtained for cephalexin synthesis. It was concluded that PGA exhibits significantly more secondary hydrolysis in comparison with AEH, unless it is reacted in a biphasic system or highly concentrated solutions (Blum and Bommarius, 2010).

   An explanation for this phenomenon is the fact that hydrolases are more active with ethers, amide being the target product. Therefore, the rate of the hydrolytic side reaction catalyzed by hydrolases is lower compared to that of hydrolysis by amidases (PGA is primarily an amidase). This increases the ratio between the synthesis and hydrolysis reaction rates for alpha-amino acid ester hydrolase (Blinkovsky and Markaryan, 1993; Zarubina et al., 2013).

2. Higher level of product accumulation in a kinetically controlled synthesis in comparison with PGA. In particular, higher rate of ampicillin synthesis and total yield of produced antibiotic can be reached. AEH form *Acetobacter turbidas* and PGA from *E. coli* were tested in synthesis of ampicillin under optimal conditions. For AEH the rate of antibiotic synthesis was much higher than the rates of ester and antibiotic hydrolyses as well as it produced much higher yields of ampicillin (85%) than PGA (36%), which also rapidly hydrolysed already synthesized ampicillin. In general, it is characteristic of PGA, that the ampicillin synthetic yields are very poor even using high excesses of acyl donors (Fernandez-Lafuente et al., 2001).

3. Higher enantioselectivity, as compared to PGA. Again, the reactions were conducted under optimal conditions: a racemic mixture of L / D-PGME was used for ampicillin preparation and the amount of L-ampicillin was measured. While PGA from *E. coli* produced a final ampicillin yield (regarding the ester transformed into antibiotic) of only 17% (maximum yield should be 50%), contaminated with a 3–4% of L-

ampicillin, the AEH from *A. turbidans* produced the yields of L-ampicillin similar to those obtained with pure enantiomer: more than 42% with enantioselectivity towards D-isomer higher than 200 (Fernandez-Lafuente et al., 2001).

4. Lower pH optimum (pH 6) compared to PGA (pH 7.5-8). The slightly acidic pH optimum of AEH is advantageous for beta-lactam stability. (Schroen et al., 1999).
5. An ability to accept charged substrates. Together with the ability of AEH to accept various beta-lactam nuclei without cleaving them, this makes alpha-amino acid ester hydrolase suitable for generating widely used antibiotics such as ampicillin and cephalexin (Blinkovsky and Markaryan, 1993).
6. Absence of inhibition by phenylacetic acid in contrast with PGA. This is due to the fact that AEH requires an alpha-amino group on the substrate, which makes an inhibition by phenylacetic acid (on of the side-products during beta-lactam synthesis) impossible (Blinkovsky and Markaryan, 1993).
7. Perspectives of use in a one-pot synthesis of ampicillin directly from penicillin G, thus eliminating the isolation step of the intermediate 6-aminopenicillanic acid (6-APA) (Blum and Bommarius, 2010).

Alfa-amino acid ester hydrolase does not necessarily need to be considered as a replacement of PGA. Several research groups describe the process of semi-synthetic beta-lactam production using a two-enzyme cascade (”one-pot” systems: a single vessel, without the purification of intermediate compounds) with both immobilized penicillin G acylase and then alpha-amino ester hydrolase. These systems significantly reduced the reaction times and outperformed “one-pot” system based solely on PGA (Blum et al., 2010; Wu et al., 2010).

Alfa-amino acid ester hydrolase as biocatalyst still suffers from several drawbacks:

1. Incapability of hydrolysis of penicillin G to yield 6-APA (Polderman-Tijmes et al., 2002).
2. Undesired primary hydrolysis of the activated acyl side chain, d-PGME, and the secondary hydrolysis of the antibiotic (Wegman et al., 2001). Although several experiments mentioned above suggest that these unwanted side reactions are negatively affecting the yield for PGA to a greater extend than for AEH.
3. Possible decrease in specificity. Presence of a hydroxyl group on the *p*-position of the phenylglycine side chain results in a drastic decrease of the specificity compared to the analogue without it (Ye et al., 2012). As a result, synthesis of amoxicillin is very low.

4. Relatively lower stability. When compared to the *E. coli* penicillin G acylase, with its optimum temperature of 49 °C and denaturation temperature of 64.5 °C (Ospina et al., 1992), the *X. campestris pv. campestris* AEH has lower kinetic thermostability. Its optimum temperature is 25 °C and an observed half-life at 30 °C is 5 minutes (Blum and Bommarius, 2010). A certain stabilization of the quaternary structure of AEH can be obtained by eliminating the phosphate ions from reaction medium (Fernandez-Lafuente et al., 2001).

In spite of numerous advantages of AEH its potential as a catalyst is not yet fully explored and the penicillin G acylase from *E. coli* remains the enzyme of choice for industrial production. While PGA activity has been detected in approximately 40 different microorganisms (including yeast, filamentous fungi, and bacteria) (Srirangan et al., 2013), about 20 different *pga* genes were cloned (as compared to only four genes previously cloned for AEH) and at least 14 structural genes were sequenced (Grulich et al., 2013), with the screening for novel penicillin acylases still on-going, the information available on the biochemistry of alpha-amino acid ester hydrolases is currently not sufficient, and further studies are required before the industrial application can be considered. Cloning of the gene that encodes AEH and its overproduction would be justified because the expression in natural hosts is relatively low.

2 Aims of the thesis

The present diploma thesis aims to:

1. Describe the microbial diversity of pharmaceutically polluted soil.

2. Identify novel alpha-amino acid ester hydrolases among cultivable bacteria from the polluted soil.

3. Construct a bacterial expression system for the novel *aeh* gene.

It should be mentioned here, that presented diploma thesis is a part of a bigger project on-going in the laboratory of enzyme technology, that includes a deeper insight into microbial diversity of pharmaceutically polluted soils and potential effects of such pollution on the rise of antibiotic resistance. Aside from functional screenings among cultivable organisms, the construction and screening of metagenomic library will be performed in the search for both novel PGA and AEH genes (for further details please refer to the discussion chapter).

3 Materials and methods

3.1 Materials

3.1.1 Microorganisms, plasmids and primers

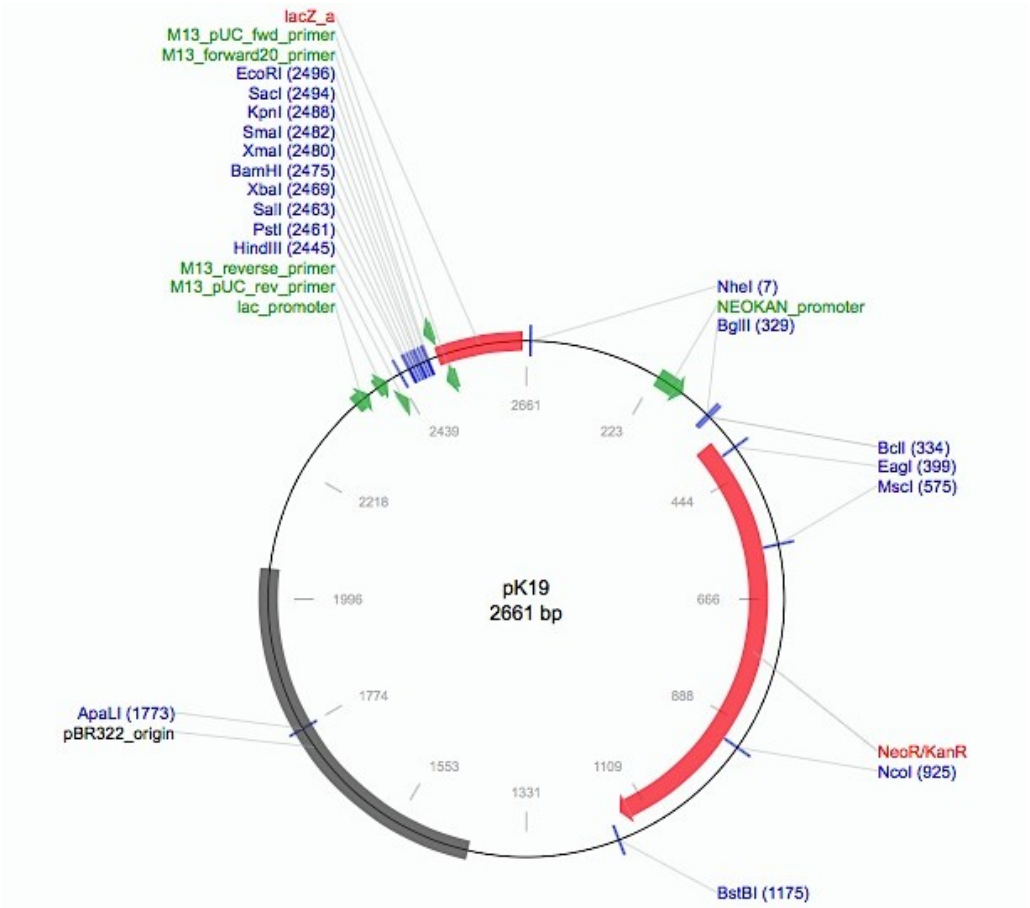
Escherichia coli BL21(DE3) – genotype: F– *omp*T *hsd*SB(rB–, mB–) *gal dcm* (DE3) (Invitrogen, USA). An *E. coli* B strain with DE3, a λ prophage carrying the T7 RNA polymerase gene. Transformed plasmids containing T7 promoter driven expression are repressed until IPTG induction of T7 RNA polymerase from a *lac* promoter.

Escherichia coli TOP10 – genotype: *hsd*R, *mcr*A, *lac*ZΔM15, *end*A1, *rec*A1 (Life Technologies, USA). Chemically competent *E. coli* cells suitable for blue-white screening.

Bacillus cereus - an endemic soil-dwelling Gram-positive rod-shaped bacterium. To this date, AEH from *Bacillus cereus*, which cloning and expression is described in this thesis, was not yet characterized.

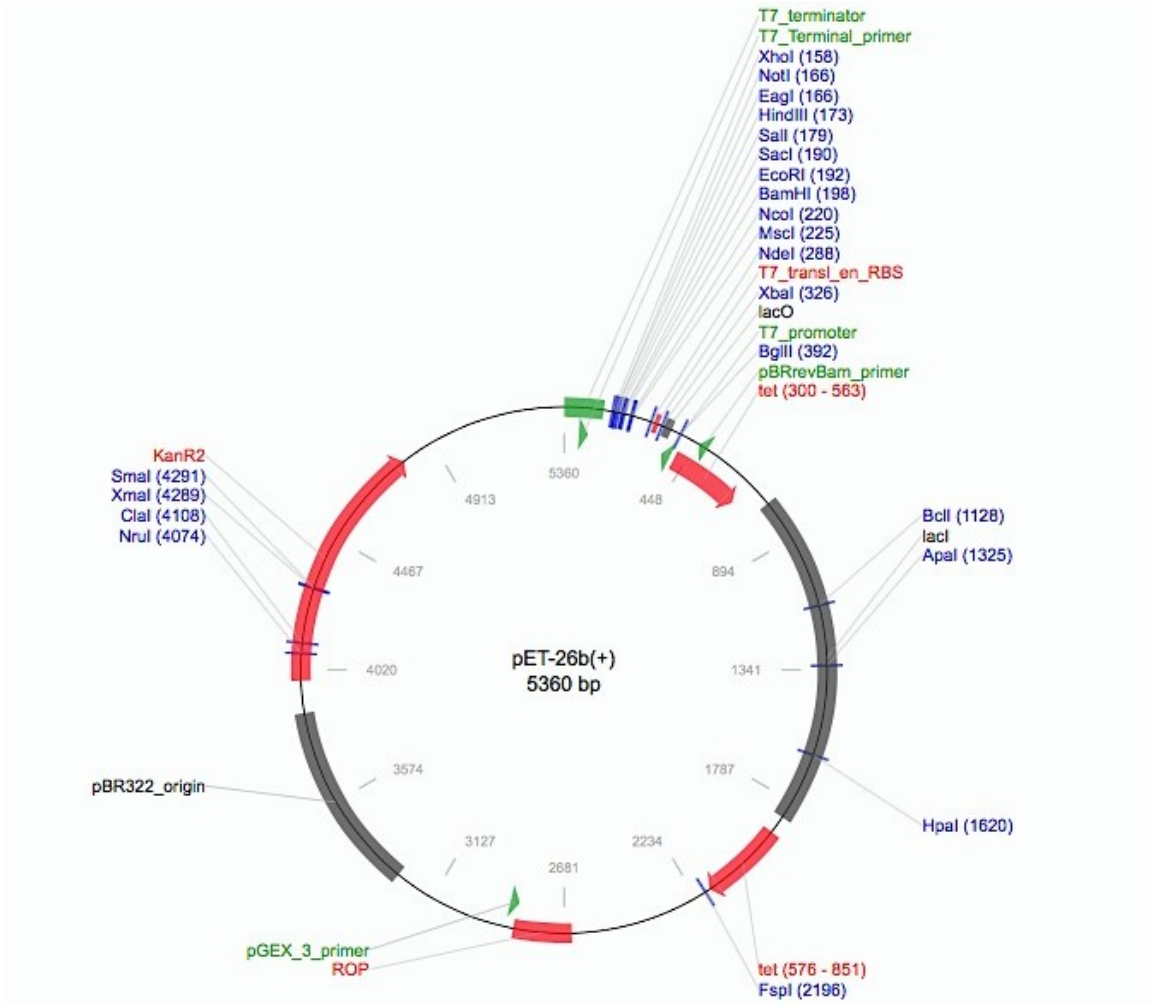
pK19 – a multi-copy kanamycin-resistance plasmid (Life Technologies, USA), containing the pUC *lac*Zα- complementation peptide and the pUC19 multiple cloning site (Pridmore, 1987).

Figure 7: A sketch showing the pk19 vector and the multiple cloning site region, created with Addgene online tool http://www.addgene.org/analyze-sequence/ (30.03.2015)



pET26b - Bacterial vector with a signal sequence and a kanamycin resistance marker for expression of proteins in the periplasm (Novagen, USA). Target genes are cloned under the control of the T7 promoter and no expression occurs until a source of T7 RNA polymerase is provided. The expression is induced by the addition of IPTG or lactose.

Figure 8: A sketch showing the pET26b vector and the multiple cloning site region, created with Addgene online tool http://www.addgene.org/analyze-sequence/ (30.03.2015)



Melting temperatures for primers were generated by the web tool OligoAnalyzer 3.1: http://eu.idtdna.com/calc/analyzer (27.03.2015). *Nde*I an *Xho*I restriction sites in specific primers are underlined. To design the degenerate primers, highly conserved motifs of AEH and COnsensus-DEgenerate Hybrid Oligonucleotide Primer (CODEHOP) strategy (Rose et al., 1998) were used.

Table 1: Specific and degenerate primers

Primer	Recognition sequence	Tm	
AEH_Fwd190	5´-GAYGCNTGGGAYACNRTNGANTGGYT-3´ For amino acid sequence DAWDTI(D/E)WL	62.5 ºC (mean)	Degenerate primers
AEH_Rev670	5´-TGNGGATTNCGRTCRTANARNGGRAA-3´ For amino acid sequence FPLYDRNPQ	60 ºC (mean)	
Fwd_AEH_Nde_ATG	5´-CTCTTCATATGCGCCGCATCGCTCCCTG-3´	66 ºC	Specific primers
Rev_AEH_Xho_Stop	5´-TAGCTCGAGTCAATGTACCGGCAGACTGAT-3´	63 ºC	

Primer concentration – 100 µM. Primer manufacturer – Metabion International AG, Germany.

3.1.2 Enzymes, chemicals, media and solutions

Enzymes and ladders

Alkaline phosphatase: FastAP Thermosensitive alkaline phosphatase 1 U/µl (Fermentas, USA)

Ligase: T4 DNA ligase 5 U/µl (Fermentas, USA)

Lysozyme: Lysozyme 50 mg/ml, in Glu-CDTA-Tris solution, 100 000 U/mg (SERVA, Germany)

Polymerase: Herculase II (*Pfu*-based DNA dependent DNA polymerase) 5 U/µl (Agilent, USA)

Proteinase: Proteinase K, 2ml > 600 mAU/ml (QIAGEN, Germany)

Restriction enzymes: *Nde*I, *Sma*I, *Xba*I, *Xho*I 10 U/ µl (Thermo Scientific, USA)

Ladders: 1 Kb DNA Ladder 500 µg/ml, Lambda DNA *Pst*I Digest 500 µg/ml (Agilent, USA)

Media and chemicals

LB medium (w/v) - 1% tryptone, 0.5% yeast extract, 1% NaCl (Sigma Aldrich, USA). To solidify the medium 15 g of agarose per litre of LB medium was added.

MYEGly medium (w/v) – 0.4% $(NH_4)_2SO_4$, 1.4% KH_2PO_4, 0.3% NaOH, 1% yeast extract, 1% (or 1.5% for production MYEGly) glycerol, pH 7.0. Before cultivation 1 ml of Trace elements solutions is added per 100 ml of MYEGly medium.

Trace elements solutions (w/v) - 2% $MgSO_4$, 0.1% $FeSO_4$, 0.5% $CaCl_2$ in distilled water

Agarose (SERVA, Germany)

dNTP Mix 50 mM, 25 mM each dNTP (Agilent, USA)

X-Gal 20mg/ml, IPTG 500mg/ml (Promega, USA)

DMSO (Agilent, USA)

$CaCl_2$, NaCl, KCl, NaOH, KOH, $(NH_4)_2SO_4$, KH_2PO_4, $MgSO_4$, $FeSO_4$, NaH_2PO_4, Na_2HPO_4, glycerol, ethanol, isopropanol, methanol, lactose, Phenol, Chloroform, isoamyl alcohol (Lach-Ner, Czech Republic)

Ethidium bromide 0.01g/ml, Kanamycin 50mg/ml, Ampicillin 50mg/ml (Sigma Aldrich, USA)

PEG 1000 (Polyethylene glycol) (Fluka, Germany)

DAB (*p*-dimethyl amino benzaldehyde) (Sigma Aldrich, USA). Stock solution prepared as 0.5% DAB (w/v) in methanol.

Buffers for plasmid purification

P1 (QIAGEN, Germany) + 1,5% lysozyme: 50 mM Tris-HCl, 10 mM EDTA, 100µg/ml ribonuclease A, 1,5% lysozyme, pH 8,0

P2 (QIAGEN, Germany): 200 mM NaOH

P3 (QIAGEN, Germany): 3,0 M KAc, pH 5,5

Other buffers and solutions

Tris-HCl buffer: 10 mM Tris-HCl, pH 8,5

TAE buffer: 40mM Tris-HCl, 20mM acetic acid, and 1mM EDTA pH 8,0. Prepared as a 50x stock solution

TE buffer: 1 mM EDTA, 10 mM Tris-HCl, pH 8,5

Na-phosphate buffer (pH 7.0): prepared by mixing 390 ml of 0.2 M monobasic sodium phosphate with 610 ml of 0.2 M dibasic sodium phosphate

Buffer O: 50 mM Tris-HCl (pH 7.5 at 37°C), 10 mM $MgCl_2$, 100 mM NaCl, 0.1 mg/ml BSA (Thermo Scientific, USA)

Tris/glycerol freezing solution: 65% glycerol (v/v), 25 mM Tris-Cl pH 7.5, 1M $MgSO_4$

6×DNA Loading Dye (Thermo Scientific, USA)

Commercial kits

High Pure PCR Cleanup Micro Kit (Roche, Switzerland)

High Pure PCR Product Purification Kit (Roche, Switzerland)

High Pure PCR Template Preparation Kit (Roche, Switzerland)

High Pure Plasmid Isolation Kit (Roche, Switzerland)

SoilMaster DNA Extraction Kit (Epicentre® (an Illumina company), USA)

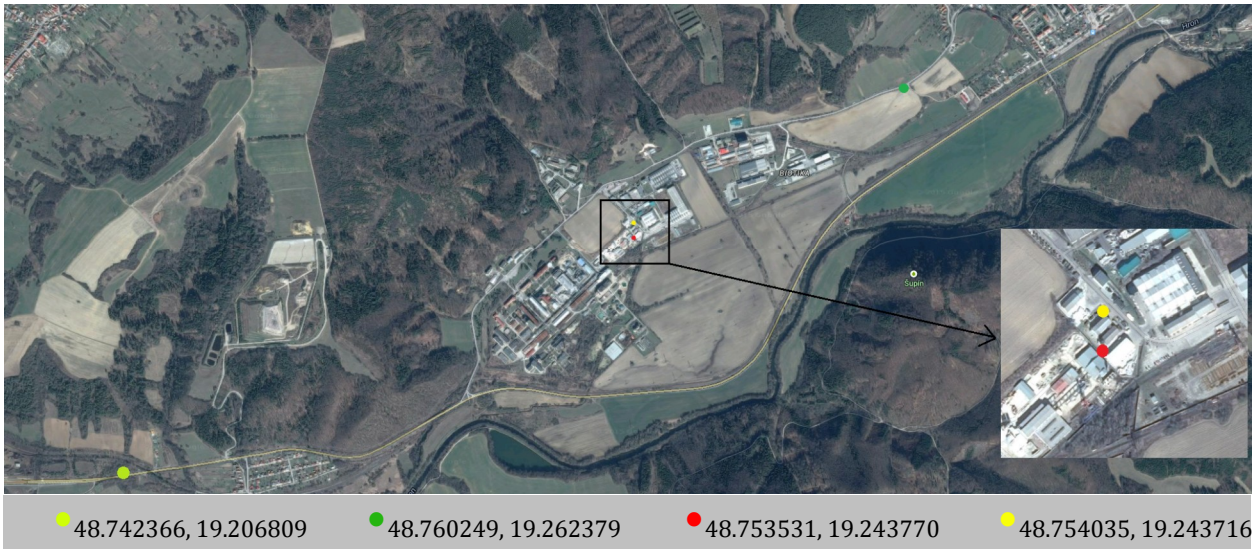
For buffer composition of Roche commercial kits please refer to the Supplement section. Buffer components of SoilMaster DNA Extraction Kit were not made public by the manufacturer.

3.2 Microbial consortia of industrially polluted soils

3.2.1 Site characterisation and sample collection

The samples were collected nearby the manufacturing plant of Biotika, a. s. in the small town of Slovenská Ľupča, Slovak Republic (48°45'133"N, 19°14'37.4"E at 371 m altitude). The factory is a producer of Penicillin G since 1956 and for the last 50 years there has been a continuous production of penicillin G, V, semi-synthetic beta-lactams and other antibiotics. Therefore it is assumed that a frequent occurrence of and higher variety of genes coding for AEH can be expected in soil microorganisms at this location.

Figure 9: Sample collection plan



The sampling took place three times: in summer of 2012, 2013 and 2014. Four cores were collected at each of the four different sampling stations: between the old and the new plant built around 10 years ago (further referred to as "polluted 1", the red dot in Figure 9), near the old plant ("polluted 2", the yellow dot), in approx. 2 km distance, to the South ("unpolluted 1", the light-green dot) and to the North ("unpolluted 2", the green dot) of the Biotika plant.

Each core was 3 cm diameter by 10 cm deep; within 48 hours they were bulked to equal one sample per station. Soils were then carefully homogenized, and for each of them pH was determined. To measure pH, soil and distilled water were mixed in 1 : 2 volumetric ratio, maintained for 10 min at room temperature to equilibrate with atmospheric CO_2, and then soil pH was estimated in the supernatant using an InoLab® pH Benchtop Meter (WTW, Germany).

3.2.2 Environmental DNA isolation

There exist two main types of procedures for isolation of environmental DNA. Direct DNA isolation is based on cell lysis within the sample matter and subsequent separation of DNA from the matter. In the indirect approach the separation of cells from the environmental material precedes the lytic release of DNA. In the present study the environmental DNA was isolated for each sample by both methods, with the products subsequently pooled together for each of the samples.

Direct approach to environmental DNA isolation

The genomic DNA was extracted using the SoilMaster DNA Extraction Kit (Epicentre, USA) according to manufacturer's protocol. Briefly, the process starts with the preparation of Spin Columns, when 550 µl of Inhibitor Removal Resin is added to each empty Spin Column to be used, followed by a centrifugation for 1 minute at 2000 x g (henceforward the centrifugation is conducted using the MiniSpin® centrifuge (Eppendorf, Germany)) to pack the column. The flow-through was decanted and the column placed in the same collection tube. The process was repeated once more and the column was moved to a clean 1.5-ml collection tube.

The lysis process starts with weighing out 100 mg of the soil sample into a 1.5 ml tube. After that 250 µl of Soil DNA Extraction Buffer and 2 µl of Proteinase K is added to the sample and the mixture is briefly mixed. To increase the yield of DNA the tube was maintained at 37°C for 10 minutes. Subsequently, 50 µl of Soil Lysis Buffer was added, and the mixture was incubated at 65°C for 10 minutes. After a 2-minute centrifugation at 1000 x g, 180 µl of the supernatant was transferred to a new tube. Then, 60 µl of Protein Precipitation Reagent was added, and the sample was thoroughly mixed by inverting the tube. This was followed by incubation on ice for 8 minutes, and 8 minutes centrifugation at 8000 x g. The supernatant (100-150 µl) was transferred directly onto the prepared Spin Column (see above) and centrifuged for 2 minutes at 2000 x g into the 1.5-ml tube. The column was then discarded.

After that, 6 µl of DNA Precipitation Solution was added to the sample, followed by incubation at room temperature for 5 minutes and centrifugation for 5 minutes at 8000 x g. The supernatant was carefully decanted. The pellet was washed with 500 µl of Pellet Wash Solution, inverted to mix, and then spin for 3 minutes at 8000 x g. The supernatant was discarded and the wash step repeated once more. Finally, the pellet was resuspended in 300 µl of TE buffer.

Indirect approach to environmental DNA isolation

The isolation of environmental DNA was executed according to a modified protocol by Sean F Brady, and represents a compilation of many methods optimized and incorporated into one procedure (Brady, 2007). First, Blending buffer of a slurry consistency was prepared. It was composed of 100 mM Tris-HCl, 100 mM Na EDTA, 1% (w/v) cetyl trimethyl ammonium bromide (CTAB) and 0.1% (w/v) SDS, pH 8.0. The Lysis buffer had the following composition: 100 mM Tris–HCl, 100 mM sodium EDTA, 1.5 M NaCl, 1% CTAB, pH 8.

After that, 125 g of soil from each sample was placed into a 500 ml Nalgene bottle. Then every sample was dispersed and homogenized in 150 ml of blending buffer. Coarse particles were collected by low-speed centrifugation at 1000 × g for 10 min at 10°C, resuspended in 150 ml of blending buffer, the cell extraction procedure was repeated two more times. Supernatants obtained during the three rounds of blending were pooled. Supernatants were centrifuged at 10 000 × g in Beckman J-68 centrifuge (Beckman Coulter, USA) for 30 min at 4°C to collect the microbial cell fraction, which was subsequently washed in 150 ml of 0.1% sodium pyrophosphate. After washing two times in 150 ml Chrombach buffer (0.33 M Tris–HCl, 1 mM EDTA, pH 8), pellets were resuspended in 8 ml Lysis buffer, with 160 µl of lysozyme (50 mg/ml), 40 µl of proteinase K added to the solution and incubated at 37°C for 30 min. Lysis was completed chemically by addition of 1 ml of 20% SDS and incubation for 2 h at 65°C in the orbital shaker (220 rpm).

Subsequently, chloroform extraction and isopropanol DNA precipitation were conducted as described below. In brief, combined and lysed cell fractions were mixed with equal volumes of the Phenol : Chloroform : Isoamyl Alcohol (25:24:1) mixture. After mixing, the sample was centrifuged and two distinct phases were formed. The aqueous phase with the desired product remained on top because the aqueous phase is less dense than the organic one. To precipitate the DNA from the recovered water phase, 0.6 volumes of isopropanol were added to the sample, followed by an overnight incubation at 4°C. The precipitates were collected by centrifugation at 16 000 × g and washed with 70% ethanol. Finally, DNA pellets were dissolved in 250 µl TE buffer. To ensure optimal quality, the sample was further purified on agarose gel and extracted using the High Pure PCR Product Purification Kit according to manufacturer's manual as described in Chapter 3.3.5.

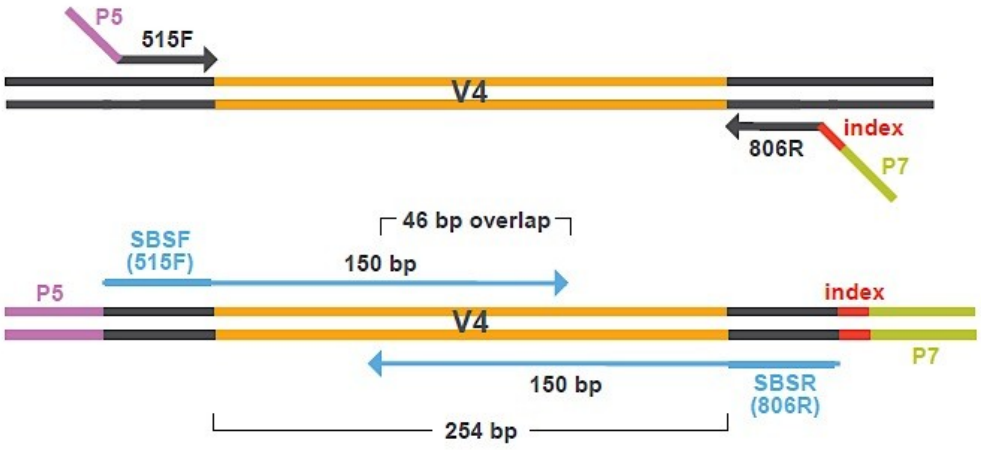
3.2.3 16S rDNA region amplification

The genes for 16S rRNA were amplified in triplicate using the primer pair:

- 515F (5'- GTGCCAGCMGCCGCGGTAA -3') concentration 10 pmol/µl and
- 806R (5'- GGACTACHVGGGTWTCTAAT -3') concentration 10 pmol/µl.

The primers included a 7-9 bp barcode for multiplexing of samples during sequencing analysis (Bates et al., 2011). This 16S primer set is designed to amplify the V4 hypervariable region (Figure 10) of Bacteria (eventually also Archaea), has few biases against specific taxa and accurately represents phylogenetic and taxonomic assignment of sequences (Liu et al., 2007).

Figure 10: Amplification strategy and paired-end reads. Source: Illumina, http://www.illumina.com/content/dam/illumina-marketing/documents/products/appnotes/appnote_miseq_16S.pdf



PCR amplifications were performed in triplicates on isolated and cleaned DNA in TPersonal thermocycler (Biometra, Germany).

Each 25–µl amplification mixture contained:

H_2O	15.7 µl
5x buffer for Herculase II polymerase	5 µl
DMSO	1.6 µl
Forward primer 515F	0.5 µl
Reverse primer 806R	0.5 µl
Environmental DNA (concentration approx. 50 ng/µl)	1 µl
Herculase II polymerase	0.2 µl

The program for PCR amplification reactions consisted of the following steps:

- initial denaturation at 94 °C, 5 min,
- 9 cycles (94 °C 90 sec, 60 °C 45 sec, 72 °C 60 sec),
- 34 cycles (94 °C 45 sec, 60 °C 45 sec, 72 °C 60 sec),
- a 10–min final extension at 72 °C.

Three parallel reactions were pooled together, purified and concentrated using the High Pure PCR Cleanup Micro Kit according to manufacturer's protocol. Briefly, total volume of each sample was adjusted to 100 µl by addition of distilled water. To each 100-µl sample 400 µl of Binding Buffer was added. The mixture was loaded into the upper reservoir of the Filter Tube that was inserted into a Collection Tube. The whole assembly was centrifuged for 45 seconds at 8000 x g and DNA was bound to a glass fiber membrane. Flow-through was discarded and the membrane was washed once with 400 µl of Wash Buffer and for the second time 300 µl of Wash Buffer. By

additional centrifugation for 1 minute at 8000 x g the excessive Wash Buffer was removed and the membrane was dried. To elute the DNA, the Filter Tube was placed into a sterile 1.5 ml tube, then 20 µl of Elution Buffer was added and the whole assembly was centrifuged at 8000 x g for 1 minute.

3.2.4 DNA library preparation

After the DNA concentration control using the Qubit® 3.0 Fluorometer (Invitrogen, Austria), the amplicons prepared from each sample, were pooled in equimolar concentrations into a single tube. The library was once again purified and concentrated to contain 1 µg of DNA in 50 µl in preparation for paired-end sequencing (2 × 150 bp) as described in the literature review section of this thesis, using the Illumina MiSeq platform at external facility GeneTiCA s.r.o. (Brno, Czech Republic).

3.2.5 Sequence processing of data from Illumina

Raw sequences were paired-end aligned using FastqJoin program. Sequencing data were filtered and trimmed using the pipeline SEED v 1.2.1 (Větrovský and Baldrian, 2013). All sequences with mismatches in tags (or primers) and ambiguous bases were removed from the dataset. All sequences shorter than 280 bp and longer than 300 bp were excluded from analysis. Chimeric sequences were detected using Uchime implementation in USEARCH v 7.0.1090 (Edgar et al., 2011) and deleted.

Subsequently, the sequences were clustered into OTUs using UPARSE implementation in USEARCH 7.0.1090 (Edgar, 2013) with a 97% similarity threshold. The consensus from each OTU was constructed from an MAFFT alignment (Katoh et al., 2009) based on the most abundant nucleotide at each position.

OTUs identification was performed by BLASTn against local database derived from Ribosomal Database Project (Cole et al., 2005) from 25.2.2014. Each OTU was assigned to the taxonomic level of genus (or nearest lower level) comparing the BLASTn best hits with the taxonomic information from the NCBI taxonomy server (http://www.ncbi.nlm.nih.gov/taxonomy). Dataset containing only bacterial sequences was used for OTU-table construction. To reduce the bias caused by different numbers of rRNA gene copies in bacterial genomes, read abundances of each OTU were divided by the copy number of the rRNA genes in the genome of the closest taxon with a complete genome sequence as described previously (Větrovský and Baldrian, 2013). To normalize for varying sequencing depths, the OTU abundance results for each sample were divided by the total number of sequences.

3.2.6 Statistical analysis

The diversity, which concerns both taxon richness and evenness of microbial community, was estimated by using the Shannon diversity index.

The Shannon diversity index is calculated as

$$H' = -\sum_{i=1}^{S}(P_i * \ln P_i)$$

where:

H' = the Shannon diversity index

P_i = fraction of the entire population made up of species i

S = numbers of species encountered

$\sum$ = sum from species 1 to species S

Differences between populations were analysed using MANOVA - multivariate analysis of variance, essentially an ANOVA (Analysis of variance) with several dependent variables - in Microsoft Excel 2013, Analysis ToolPak (normality of data was verified in advance). The input to statistical analysis was comprised of relative abundances of phyla and genera. The genera with relative abundance of at least 0.5% (if not specified otherwise) were included into analysis. Differences between groups were declared significant at $p \leq 0.05$.

Principal component analysis (PCA) was applied to identify the community structure differences in different locations, with the use of PAST software, version 3.06. The analysis is based on reducing the dimensionality of the data by identifying principal "axes" of discrimination among communities, which represent distinct combinations of microorganism taxa. Additionally, nonmetric multidimensional scaling using the Bray-Curtis dissimilarity index (PAST software, v3.06) was used to determine the patterns of community composition and pH soil parameters.

3.3 AEH ”gene mining” and construction of bacterial expression system

3.3.1 Functional screening among cultivable microorganisms

First, 1 gram of soil was placed into a sterile 15 ml plastic tube containing 10 ml of phosphate buffer (pH 7.0). This mixture was homogenized for 30 seconds by shaking at low speed on IKA® MS1 Shaker (IKA, USA; hereafter this minishaker is utilized for mixing all the samples if not stated otherwise). After homogenization, the soil suspension was serially diluted from 10^{-1} up to 10^{-7}. Then, 100 µl aliquots of serially diluted soil suspensions were plated on LB agar. The plates were incubated at 28 °C for 48 hours.

The functional screening among cultivable organisms was carried out using the chromogenic activity assays. Each colony was incubated overnight at 28 °C in a microtiter plate in LB medium. Subsequently, 50 µl of each culture was transferred to a new microtiter plate, 50 µl of 0.1 M phosphate buffer (pH 7.0) and 100 µl of 1% ampicillin solution was added to them. The cultures were incubated for 4-5 hours at 28 °C with constant shaking on Kuhner orbital Lab-Shaker (Kuhner, Switzerland) at 200 rpm, 50 mm shaking diameter. After that, the cultures were supplied with 50 µl of stop solution (prepared as a mixture of 20% acetic acid and 0.05M NaOH), and after 15 minutes with 50 µl of DAB (0.5% (w/v) *p*-dimethyl amino benzaldehyde in methanol). Because of the turbidity, the activity of each extract was judged either positive or negative against the controls by eye.

3.3.2 Chromosomal DNA purification

To isolate DNA from bacterial culture, 10 ml of LB media was inoculated with an individual bacterial colony and cultivated overnight in 28 °C with constant shaking in Gallenkamp orbital incubator (Gallenkamp, UK; orbit diameter: 32 mm; henceforward this orbital incubator is used for liquid culture cultivations in 37 °C (or 28 °C, when specified) at 200 rpm).

DNA extraction from bacterial cells was performed using High Pure PCR Template Preparation Kit (Roche, Switzerland) following the manufacturer’s protocol. First, the culture medium containing approximately 10^9 cells was centrifuged for 5 minutes at 3000 x g and supernatant was discarded, then the cells were resuspended in 200 µl of P1 (+ 1,5% lysozyme) buffer. After 15 minutes of incubation at 37 °C, the resuspended cells were transferred to a nuclease free 1.5 ml microcentrifuge tube. Subsequently 200 µl of Binding Buffer and 40 µl of reconstituted Proteinase K were added. The contents of the tube was mixed immediately and incubated at 70°C for 10 min. After that, 100 µl of isopropanol was added to the mixture.

The entire sample mixture was pipetted into the upper reservoir of High Filter Tube inserted into the Collection Tube. After the centrifugation of the entire assembly in a minispin centrifuge for 1 minute at 8000 x g, the Filter Tube was removed from the Collection Tube and the flowthrough liquid (and the Collection Tube) was discarded. Then the Filter Tube was combined with a new

Collection Tube and 500 µl of Inhibitor Removal Buffer was added to the upper reservoir of the Filter Tube. The entire assembly was centrifuged for 1 minute at 8000 × g. The flowthrough liquid and the Collection Tube were again discarded, then the upper reservoir of the Filter Tube was washed twice with 500 µl of Wash Buffer (using the centrifugation at 8000 × g). After discarding the flowthrough liquid and performing an additional 10-second centrifugation at full speed, the Filter Tube was inserted into a clean, sterile 1.5 ml microfuge tube. To elute the DNA 200 µl of prewarmed (70°C) Elution Buffer was pipetted to the upper reservoir of the Filter Tube, followed buy a centrifugation of the tube assembly for 1 minute at 8000 × g.

3.3.3 PCR amplification

PCR amplifications were performed on isolated and cleaned DNA in TPersonal thermocycler (Biometra, Germany)

Each 25–µl reaction mixture for amplification of approximately 1400 bp fragment of *aeh* gene using degenerate primers contained:

H_2O	14.9 µl
5x buffer for Herculase II polymerase	5 µl
DMSO	1.6 µl
Forward primer AEH_Fwd190	1 µl
Reverse primer AEH_Rev670	1 µl
Chromosomal DNA (50 ng/µl)	1 µl
Herculase II polymerase	0.5 µl

The program for these PCR amplification reactions consisted of:

- initial denaturation at 95 °C, 5 min,
- 35 cycles (95 °C 45 sec, 55 °C 55 sec, 72 °C 90 sec),
- a 5–min final extension at 72 °C.

The reaction mixture for amplification the complete *aeh* gene using specific primers:

H_2O	15.9 µl
5x buffer for Herculase II polymerase	5 µl
DMSO	1.6 µl
Forward primer Fwd_AEH_Nde_ATG	0.5 µl
Reverse primer Rev_AEH_Xho_Stop	0.5 µl
Chromosomal DNA (50 ng/µl)	1 µl
Herculase II polymerase	0.5 µl

The program for these PCR amplifications was set as follows:

- initial denaturation at 95 °C, 5 min,
- 35 cycles (94 °C 30 sec, 57 °C 40 sec, 72 °C 50 sec),
- a 5–min final extension at 72 °C.

3.3.4 DNA electrophoresis

Products of PCR were visualized on agarose gel electrophoresis. Agarose gel (1%) was prepared from 0.35 g of agarose and 35 ml of TAE buffer. Agarose with TAE buffer was boiled until agarose was completely dissolved and after cooling down it was poured into the electrophoretic device. 10 µl of 1% ethidium bromide was added before gel solidification. After placing a comb into tray with a gel, it was left to cool down. The solid gel was placed into electrophoretic tank Shelton Scientific IBI® Horizontal Gel Systems JSB-30 (IBI, USA) with 1 x TAE buffer. 1-5 µl of the sample with added to 1-3 µl of 6x DNA Loading Dye was loaded on gel. Electrophoresis ran at 5 V per 1 cm of gel (max 500 mA) for 40 minutes and was viewed under UV in transilluminator and photographed.

3.3.5 DNA extraction from agarose gel

DNA extraction from an agarose gel was performed using High Pure PCR Product Purification Kit according to manufacturer's manual. Briefly, desired DNA fragment was visualized by UV lamp, excised from an agarose gel using an ethanol-cleaned scalpel and transferred to a 2 ml clean tube. The weight of the gel slice was determined and 300 µl of Binding Buffer was added for each 100 mg of gel. Samples were incubated at 56 °C for approximately 10 minutes (until the gel was completely dissolved) and, after addition of 150 µl of isopropanol for every 100 mg agarose gel slice, the mixture was loaded into the upper reservoir of the Filter Tube and placed into a Collection Tube. The Collection Tube was centrifuged for 45 seconds at 8000 x g and DNA was bound to a glass fiber membrane. Flow-through was discarded and the membrane was washed with 500 µl of Wash

Buffer twice. By additional centrifugation for 1 minute at 8000 x g the excessive Wash Buffer was removed and the membrane was dried. To elute the DNA, the Filter Tube was placed into a sterile 1.5 ml tube, then 30 µl of Elution Buffer was added and the whole assembly was centrifuged at 8000 x g for 1 minute.

3.3.6 Plasmid and fragment restriction

In separate tubes the plasmid vector and the DNA fragments were digested with the appropriate restriction enzymes.

In a microfuge tube the following mixture was set up for each digesting reaction:

H_2O	14 µl
10x restriction enzyme buffer	2 µl
DNA (1-2 µg)	3 µl
Restriction enzyme(s)	1 µl

For pK19 vector restriction enzyme *Sma*I was used (to generate blunt ends).

PCR product of the complete *aeh* gene and pET26b vector plasmids were digested with the same restriction enzymes (0.5 µl of *Nde*I and 0.5 µl of *Xho*I).

The reaction was performed at 37°C for 1-2 hours. After the incubation time was over, the vector ends were defosforylated with alkaline phosphatase to avoid self-ligation in the next (ligation) step. Therefore 1 µl of alkaline phosphatase was added to the linearized plasmid and incubated for another 30 minutes at 37°C. Sample quality was then checked by electrophoresis in agarose gel, desired products were excised and purified as described in Chapter 3.3.5.

3.3.7 DNA ligation

In a microfuge tube, the following 20-µl ligation mixture was set up:

5x T4 DNA ligase buffer	4 µl
PEG	2 µl
Plasmid vector (10 ng/µl)	3 µl
DNA fragment (16 ng/µl)	10 µl
T4 DNA ligase	1 µl

The ligation mixture was maintained overnight at 22°C

3.3.8 Transformation of DNA into competent cells and selection of transformants

Competent bacteria (Top 10 or BL21 (DE3)) stored in suspension in - 80°C were slowly thawed on ice. 100 µl of bacterial suspension per sample were supplied with 20 µl of ligation mixture, gently mixed (without the help of minishaker) and incubated for 30 minutes on ice. Subsequently, the cells were exposed to a heat shock - 42°C for 90 seconds and immediately placed on ice for 1.5 minutes. After that 400 µl of LB medium was pipetted to the sample and the mixture was incubated for 30 minutes at 37°C under constant mixing. The bacterial suspension spread out on the surface of LB agar supplied with kanamycin (200 µl of suspension on each plate). Bacteria on the plates were cultivated at 37°C overnight bottom side up.

The successfulness of transformation with vector pK19 could be verified by the blue-white screen, a well-known method based on the fact that X-gal is converted by beta-galactosidase into an insoluble dense blue compound. Blue colonies will contain plasmid vectors without the insert, whereas white colonies contain plasmid vectors with the insert disrupting the gene for the alpha-peptide of beta-galactosidase. To apply this method, before solidification LB agar was supplied with 60 µl of kanamycin solution, 200 µl of of X-gal and 20 µl of IPTG for each 100 ml of LB agar medium. The presence of insert DNA in the cloned pET26b constructs was validated by colony PCR and sequencing.

3.3.9 Colony PCR

Using a micropipette tip a small amount of a single putative colony was taken off the plate. The tip was inserted into the PCR mixture (see below) and pipetted up and down.

PCR mixture contained:

H_2O	16.9 µl
5x buffer for Herculase II polymerase	5 µl
DMSO	1.6 µl
Forward primer FW_AEH_Nde_ATG	0.5 µl
Reverse primer Rev_AEH_Xho_Stop	0.5 µl
Herculase II polymerase	0.5 µl

The program for these PCR amplifications was set as described in the last part of Chapter 3.3.3. Individual transformants were lysed during the initial heating step, which caused the release of the plasmid DNA from the cell, so it could serve as a template for the amplification reaction. Primers designed to specifically target the insert DNA and subsequent visualization on gel electrophoresis were used to determine if the construct contains the desired *aeh* gene.

3.3.10 Plasmid DNA purification

For small scale purifications (e.g. for plasmid verification after bacterial transformation), 2 ml of LB media (with kanamycin) was inoculated with an individual bacterial colony of successfully transformed cells and cultivated overnight in 37°C with constant shaking (Gallenkamp orbital incubator). Plasmid isolation was performed according to a modified protocol described by Birnboim. "The principle of the method is selective alkaline denaturation of high molecular weight chromosomal DNA while covalently closed circular DNA remains double-stranded" (Bimboim and Doly, 1979).

The overnight culture was transferred to 2 ml Eppendorf tube and pelleted by centrifugation at 4,500 x g for 5 minutes (Eppendorf mini spin centrifuge). Supernatant was discarded and pellet was resuspended (IKA-vortex minishaker) in 200 µl of buffer P1 (+ 1,5% lysozyme) and incubated for 15 minutes at 37°C. Subsequently 400 µl of buffer P2 was added to the mixture and the sample was gently mixed by inverting the tube 5 times. The mixture was maintained at room temperature for 5 minutes. After that 300 µl of buffer P3 was added, the sample was again gently mixed by inversion and incubated for 15 minutes at 0 °C. This was followed by a centrifugation (12000 x g, 3 minutes). Subsequently 600 µl of supernatant was transferred to a new 2 ml Eppendorf tube, followed by an addition of 200 µl of ice-cold ethanol. The samples were incubated at -20°C for 30 minutes, then the pellet was collected by centrifugation (12000 x g, 3 minutes) and dissolved in 25 µl of Tris-HCl buffer.

For large-scale purifications (e.g. for subsequent transformation), again 2 ml of LB media (with kanamycin) was inoculated with an individual bacterial colony of successfully transformed cells and cultivated overnight in 37°C with constant shaking. Then, plasmid isolation was carried out using High Pure Plasmid Isolation Kit (Roche, Switzerland).

Briefly, the overnight culture was pelleted by centrifugation at 4,500 x g for 10 minutes at 4°C. Supernatant was discarded and pellet was resuspended in 250 µl of Suspension Buffer + RNase. The resuspended bacterial pellet was treated with 250 µl of Lysis Buffer, mixed gently by inverting the tube 5 times and incubated for 5 minutes at room temperature. Subsequently, the lysed solution was treated with 350 µl of chilled Binding Buffer, mixed by inverting the tube 5 times and maintained on ice for 5 min. The cloudy solution was centrifuged for 10 minutes at 8000 × g.

The entire supernatant was transferred into upper buffer reservoir of the Filter Tube and the entire High Pure Tube assembly (High Pure Filter Tube inserted into a Collection Tube) was centrifuged for 1 minute at 8000 × g (minispin). After centrifugation the flowthrough liquid was discarded, the Filter Tube reinserted in the same Collection Tube and the pellet was washed first with 500 µl of Wash Buffer I, then with 500 µl of Wash Buffer II. The plasmid DNA was eluted into a sterile 1.5 ml microcentrifuge tube by adding of 50 µl of Elution Buffer to the upper reservoir of the Filter Tube and centrifuging the tube assembly for 1 minute at full speed 8000 × g (minispin).

3.3.11 DNA Sequencing and sequence analysis

The samples were sent for sequencing to an external facility IMB AS CR. DNA was sequenced with BigDyeTM Terminator Cycle Sequencing Kit and the genetic analyser ABI PRISM 3130x1 (Applied biosystems). Individual sequences were then manually edited using the program BioEdit version 7.2.5 and compared to those collected in GenBank database (http://www.ncbi.nlm.nih.gov/genbank/).

3.3.12 Cultures in continuously stirred tank reactor

All solutions of MYEGly medium components were prepared separately and sterilized for 30 minutes at 121 °C. Culture inoculum preparation: a Working Cell Bank (WCB) and flask cultures of 1st and 2nd generations were prepared.

To prepare the WCB, the culture was grown in 100 ml of LB medium with kanamycin for 16-18 hours at 28 °C with constant shaking at 200 rpm in orbital incubator. After that 0.6 ml of grown culture was mixed with 0,6 ml of Tris/glycerol solution in cryovial and the mixture was immediately placed to -80 °C.

To prepare the flask culture of 1st generation, one loop of the WCB was used to inoculate a petri plate with solid LB medium and kanamycin. The plates were cultivated at 28 °C for 24 hours. After that two loops from solid LB medium were transferred to 100 ml of liquid LB medium with kanamycin in 500 ml flask. The cultivation of inoculum of first generation was carried out on shaker for 18 ours at a temperature of 28 °C (OD_{600} 4-5, pH 8.1).

To prepare the flask culture of the 2nd generation, 2 ml of first generation inoculum was added to 100 ml of MYEGly medium in 500 ml flask. Trace elements solutions were added into the medium, in amount of 1 ml of each element per 100 ml of medium. Cultures were grown for 10 hours (OD_{600} about 6-8, pH 6-7) at 28 °C in orbital incubator.

Subsequently, 150 gram of second-generation inoculum was transferred to a bioreactor Biostat MD (B.Braun Biotech, Germany) with 6 liters of sterilized MYEGly medium and trace elements. The conditions of cultivation were as follows: temperature - 28 °C, pH was maintained at 7.0 with 25% solution of NH_4OH, aeration rate – 6 NL of air per minute, foam control with 10% antifoam suspension. The value of pO_2 was controlled by stirring in the range of 300-900 rpm. The growth of the culture was not limited by availability of oxygen: level of pO_2 was maintained at the value of 20% of fully saturated medium by air under given culture conditions..

Glycerol feeding starts (40% (w/v) glycerol is used as a carbon source solution) approximately after 10 hours of cultivation, when the upper limit of stirring speed (800-900 rpm) was achieved. There are two lactose induction doses: 150 grams of 10% lactose at OD_{600} 40-45 (approx. after 15-16 hours), and then 75 grams of 10% lactose at OD_{600} 50-60 (approx. after 22-23

hours). Stringent sterility controls of the fermentation medium and controls of culture purity by cultivation on solid LB medium were conducted. The culture in bioreactor was sampled at regular intervals and OD_{600}, CDW and activity of AEH were assayed.

3.3.13 Activity assays

Activity of AEH was measured for the prepared bacterial expression system (flask and fed-batch cultivations) and the original *B. cereus* strain (shaken flask cultivation only) isolated from soil. The measurements were performed by a kinetic assay using the Halo DB-20 Spectrophotometer (Dynamica, UK). One unit of activity (U) is defined as the amount of the enzyme that catalyses the conversion of 1 μmol of substrate per minute at 37 °C.

Preparation of the reaction mixture from stock solutions

A: 0.1 M Na-phosphate buffer (pH 7.0)

B: 1% (w/v) solution of ampicillin in buffer A (pH 7.0, prepared fresh)

C: Acetic acid - 20% solution (100 ml of glacial acetic acid up to 500 ml with distilled water)

D: NaOH 0.05M solution: 1 g of NaOH in 500 ml of distilled water

E: mixture of solutions C : D = 2 : 1 (to stop reaction)

F: p-dimethyl amino benzaldehyde (DAB). Stock solution preparation: 0.5 g DAB in 100 ml of methanol (kept in dark at + 4 ^{O}C).

G: Standard solution of 6-APA (5 mM solution): 108 mg of 6-APA dissolved in buffer A and volume adjusted up to 100 mL (54mg/50mL)

Preparation of samples for activity assays

The samples of flask cultures for measuring activities were prepared by an overnight cultivation in 10 ml of MYEGly medium and 0.2 ml of 10% (w/v) lactose as an inducer at 28 °C with constant shaking. The samples coming from fed-batch (fermentation) cultivation were collected at regular intervals. Subsequently 8 ml of cell culture was centrifuged (8000 x g for 5 min.) in calibrated test tubes. Supernatant was discarded and the pellet was frozen two times for at least 1 h (-24 °C). The pellet was resuspended in 4 ml of buffer A (a half of the original culture volume). Resuspended pellet was 10 times diluted and subsequently sonicated with Sonicator 3000 (Misonix, USA) with the intensity 14 W per 5 ml, amplitude 2.5 and 0 °C temperature. Afterwards, 0.25 ml of buffer A was added to 1 ml of sonicated sample solution and the mixture was pre-warmed for (5 minutes at 37 ^{O}C in a water bath Tempette TE-8D (Techne, UK)). Additionally, 1.25 ml of buffer A was prepared as a blank.

Enzymatic reaction was initiated (t=0) by addition of 1.25 ml of solution B (substrate) to 1.25 ml of pre-warmed sonicated sample solution (as above). The reaction mixture was sampled (0.5 ml) at 3th, 6th, 9th, and 12th min and the reaction was immediately stopped by addition of 3 ml of solution E to the sample. The samples were then centrifuged for 3 minutes at 8000 x g, the supernatants were transferred into clean test tubes and 0.5 ml of solution F (DAB) was added to them. After 15 minutes absorbance was measured at 415 nm (A_{415}).

Calibration curve for DAB: the solution G was diluted with buffer A to prepare 6-APA solutions of 1.0 - 5.0 µmol/ml. After plotting A_{415} (X axis) against concentration of 6-APA (Y axis) factor f_{DAB} was calculated (f_{DAB} = 7,3909).

Calculation of the volumetric activity (VA):

VA (U/L) = f * 2.5 * 1000 (multiplied/divided if diluted/concentrated)

where:

f	slope of a straight line, obtained by plotting time of sampling (min) on X axis against (A_{415} * f_{DAB})on Y axis
2.5	sample dilution in assay
1000	re-calculation to volume of 1 l

To calculate the specific activity (SA, in U/g_{cdw}), the volumetric activity was divided by the measured cell dry weight.

3.3.14 Determination of dry mass and assay of OD_{600}

To determine the bacterial dry mass (or cell dry weight (CDW) in g/l), measured volumes of the culture samples where centrifuged, the biomass was washed twice with distilled water, placed on preweighed membrane filters and dried at 105 °C for approximately 2 hours (until the mass was no longer changing) in the oven.

Growth curve of a culture in a stirred tank was constructed by determining the optical density (a measure of the light scattered by the bacterial suspension) at 600 nm (OD_{600}) using the Halo DB-20 Spectrophotometer (Dynamica, UK).

4 Results

4.1 Microbium consortia characterization

In this part of the study a high-throughput sequencing of the V4 variable region of the 16S ribosomal RNA gene was used to test the hypothesis that bacterial communities were altered under the pressure of pollution with antibiotics.

Samples from polluted and unpolluted soils were collected and homogenized as described previously. Soil pH was measured for each sample: an average value was 7.2, with minor insignificant differences among different sampling sites.

To determine the composition and diversity of soil communities, genomic DNA was extracted from each of the 12 samples using two different methods as described in Chapter 3.2.2. The whole extraction procedure was repeated one more time for each soil sample, thus creating a second independently treated set of samples (replicate). For each of the 24 samples, V4 variance region of 16S rDNA was amplified in triplicates using specific primers containing 7-9 bp barcode for multiplexing of samples during sequencing analysis. Three parallel reactions were pooled together, purified and concentrated using the High Pure PCR Cleanup Micro Kit.

The amplified DNA concentration was verified using the Fluorometer (see Table 2), the amplicons prepared for each sample, were pooled in equimolar concentrations into a single tube. After that, the library was once again purified, concentrated send for paired-end sequencing using the Illumina MiSeq platform to an external facility GeneTiCA s.r.o.

Table 2: Amplified DNA concentration of two independent sets of samples

Samples (1st replicate)	Forward primer	Reverse primer	Concen-tration (ng/µl)	Samples (2nd replicate)	Forward primer	Reverse primer	Concen-tration (ng/µl)
Polluted_2 2012	515F_T001	806R_T003	15.70	Polluted_2 2012	515F_T048	806R_T054	19.80
Polluted_1 2012	515F_T001	806R_T007	15.80	Polluted_1 2012	515F_T048	806R_T101	30.60
Unpolluted_1 2012	515F_T001	806R_T011	9.46	Unpolluted_1 2012	515F_T048	806R_T105	13.40
Unpolluted_2 2012	515F_T001	806R_T050	16.50	Unpolluted_2 2012	515F_T048	806R_T107	10.20
Polluted_2 2013	515F_T001	806R_T052	11.10	Polluted_2 2013	515F_T048	806R_T111	8.58
Polluted_1 2013	515F_T001	806R_T054	17.30	Polluted_1 2013	515F_T096	806R_T003	12.40
Unpolluted_1 2013	515F_T001	806R_T101	8.58	Unpolluted_1 2013	515F_T096	806R_T007	12.20
Unpolluted_2 2013	515F_T001	806R_T105	10.40	Unpolluted_2 2013	515F_T096	806R_T011	9.64
Polluted_2 2014	515F_T001	806R_T107	8.42	Polluted_2 2014	515F_T096	806R_T050	8.32
Polluted_1 2014	515F_T001	806R_T111	8.38	Polluted_1 2014	515F_T096	806R_T052	11.10
Unpolluted_1 2014	515F_T048	806R_T007	8.06	Unpolluted_1 2014	515F_T096	806R_T101	7.50
Unpolluted_2 2014	515F_T048	806R_T011	28.60	Unpolluted_2 2014	515F_T096	806R_T105	27.00

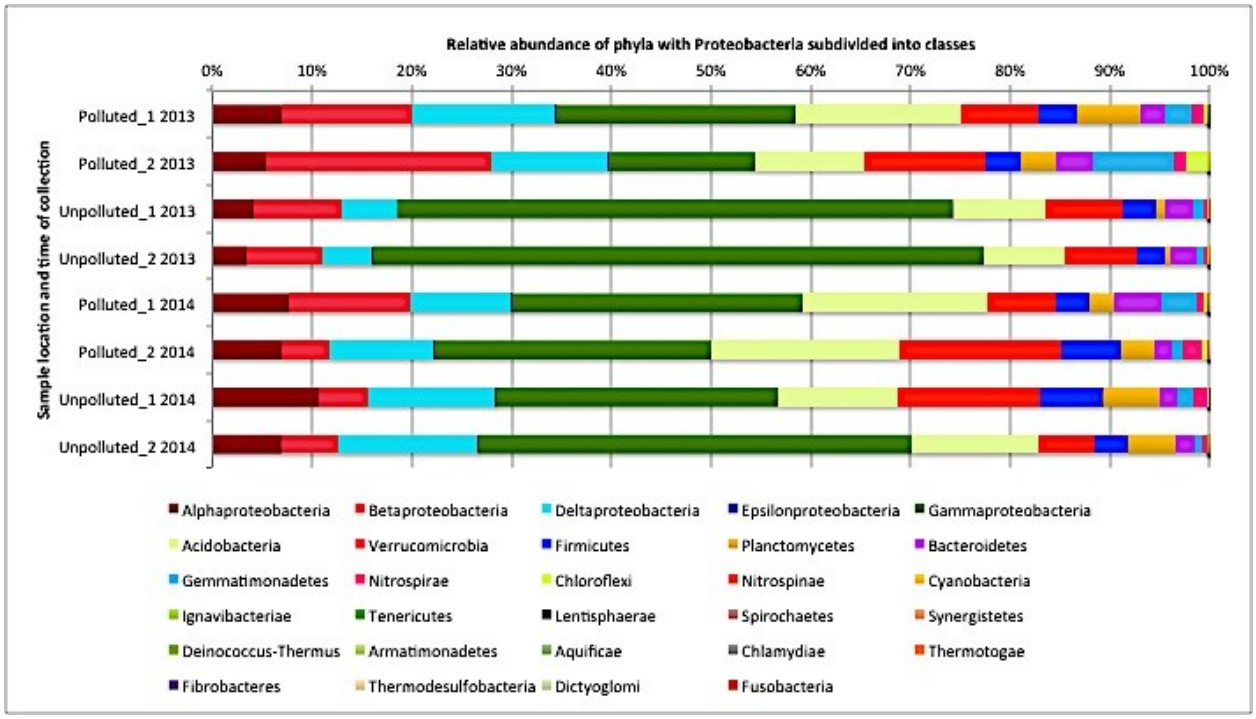
Illumina-based sequencing of 16S rRNA amplicons produced a total of 433433 paired-end joined sequences that passed quality filtering. Retrieved raw sequences were processed as described in Chapter 3.2.5. After all the sequences shorter than 280 bp and longer than 300 bp, with mismatches in tags (or primers) and ambiguous bases were removed from the dataset, there remained 362614 sequences, in average 13947 sequences per sample. Subsequently, the sequences were clustered into OTUs with a 97% similarity threshold and OTUs identification was performed by BLASTn against local database derived from Ribosomal Database Project, where each OTU was assigned to the taxonomic level of genus comparing the BLASTn best hits with the taxonomic information from the NCBI taxonomy server.

Although the present study was designed to determine the bacterial diversity, utilized primer set additionally amplified a certain number of sequences assigned to Archaea (18138 sequences) and organelles (10897 sequences). For 48 sequences the taxonomic identification failed, whereas the remaining set of bacterial 16S rDNA sequences constituted for 10092 OTUs.

To simplify the investigation and prevent from potential biases connected with a long-term storage of samples that were originally collected for the purposes of screening among cultivable bacteria, the decision was taken to exclude the samples collected in 2012, and continue with the analyses for the samples from 2013 and 2014. The data set was normalized, corrected as to the copy number of the rRNA genes, and relative abundances of OTUs, genera or phyla made up an input for statistical analyses. First, the differences among samples from unpolluted soil collected at different sampling sites and different time periods were investigated. No significant difference was revealed by MANOVA among all unpolluted samples (collected in both time periods) including replicates ($p = 0.75$). The same is true for the group of samples collected at different time periods at polluted 1 and polluted 2 sites ($p = 0.16$). Additionally, no significant difference in community structure was found between corresponding samples from two independent DNA isolation replicates. Further statistical analysis by MANOVA declared significant difference at the genus level between samples from polluted and unpolluted sites collected in 2013 ($p = 0.05$), but no significant difference for the same samples collected in 2014 ($p = 0.86$). The most significant resolution was achieved when polluted samples from both time periods were compared to the unpolluted sample from 2013 ($p = 0.01$). Phylum-level MANOVA demonstrated that bacteria from polluted and unpolluted areas share similar phyla diversity, and community changes are not reflected at this level.

In total 26 different bacterial phyla consisting of 918 genera were detected across all soil samples. The distribution of sequences among all identified phyla and most abundant genera is depicted in Figure 11 and Figure 12.

Figure 11: Distribution of sequences belonging to individual phyla (Proteobacteria further subdivided into classes) identified at sampling sites, with two replicates merged for each site.
The phyla and respective legend are shown in descending order from left to right.
Created in Microsoft Excel 2011.

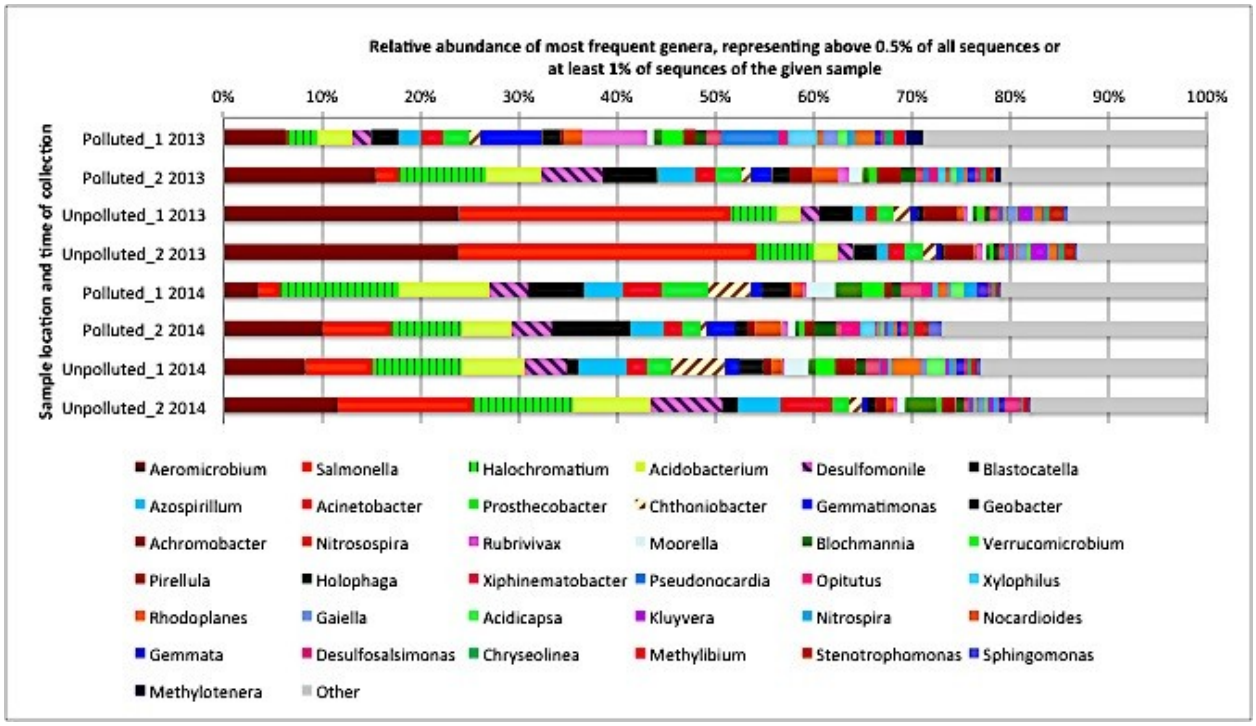


The bacterial community composition of all samples at phylum level was relatively uneven, with 4 main phyla (consistent among samples) constituting above 80% of abundance in each sample. The most abundant were Proteobacteria (50.9% of sequences on average), followed by Actinobacteria (17.6%), Acidobacteria (11.2%) and Verrucomicrobia (8.1%).

Within Proteobacteria, the majority of sequences were assigned to the classes of Gammaproteobacteria (mean value 28.6%) and Deltaproteobacteria (8.8%). Betaproteobacteria (8% mean value), were especially abundant in polluted samples collected in 2013 (14.1%), whereas Gammaproteobacteria in this group of samples represented only 15.5% of sequences, as contrasted to the above-mentioned average of 28.6%. On the contrary, an unpolluted sample collected in 2013 contained 41.9% of sequences belonging to Gammaproteobacteria.

Among other abundant phyla, Firmicutes (mean value 3.43%), Planctomycetes (2.90%), Bacteroidetes (2.24%), Gemmatimonadetes (1.97%), Nitrospirae (0.79%) and Chloroflexi (0.43%) should be mentioned.

Figure 12: Distribution of the most abundant genera (representing over 0.5% of all sequences) identified at sampling sites, with two replicates merged for each site.
The genera and respective legend are shown in descending order from left to right.
Created in Microsoft Excel 2011.

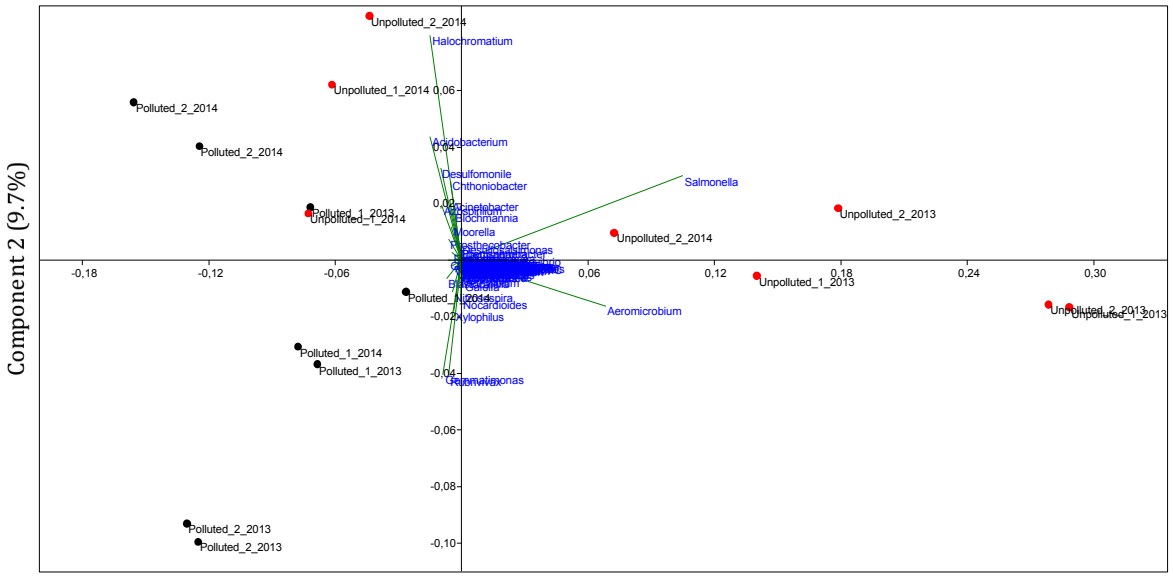


At genus level (Figure 12), the relative abundance of dominant genera was significantly dissimilar in different samples. *Aeromicrobium*, with 12.9% on average is the most abundantly represented genus in all samples excluding unpolluted ones collected in 2013, where *Salmonella* dominated with the abundance of 29% of all sequences in these samples. Bacteria belonging to *Halochromatium* genus appear to be more populous in 2014 (9.7%) than in 2013 (5.6%). *Acidobacterium* strains seems to be equally represented in polluted (7.2%) and unpolluted (7.1%) samples collected in 2014, in contrast with 2013, when this genus was more abundant in polluted (4.6%) than unpolluted samples (2.5%). On the other hand, *Blastocatella* abundance appears to be higher in polluted samples (mean value 5.5%) than in unpolluted (2.1%) in both time periods; the same is true for the *Gemmatimonas* genus.

The bacterial community at this level was observed to be highly uneven: 150 taxa (16% of total) represent 95.6% of all retrieved sequences. Additionally, the community composition of unpolluted samples collected in 2014 appears to be transitive between other polluted and unpolluted samples, which is in compliance with the results of principal component analysis.

Principal component analysis was used to identify community structure differences for individual samle sites. It showed that the two types of communities (polluted and unpolluted) could be well separated at the genus level (see Figure 13), and also at phylum level, however, with lower resolution (data not shown).

Figure 13: Principal component analysis of bacterial communities from the samples collected in 2013 and 2014, including separated replicate samples. A total of 15 axes were necessary to explain all variation. Samples from the polluted locations are labelled with black dots, from unpolluted – with red dots. The analysis is performed on 100 most abundant genera. The genera with the highest influence on distribution are denoted in blue. Created in PAST v3.06.



In the genus-level PCA, a total of 15 axes were necessary to explain all variation. The first two axes explained 89% of total variability (Figure 13). The analysis demonstrates that the replicates of the same sample are clustered and the communities dwelling in polluted (marked with black dots) and unpolluted (red dots) areas can be well separated, although there is no clear boundary, as the unpolluted samples collected in 2014 (especially from site 1) approach the polluted ones.

To gain an overview of the biodiversity of the studied bacterial community, the Shannon index was calculated. Separate analyses were performed at the genus and phylum levels, and the results were aggregated across the times of sample collection and replicates. Statistical analysis revealed a significant difference (p = 0.05) between the polluted and unpolluted sample sites, and the average genus-level Shannon diversity indices were 4.10 for polluted area and 3.45 for unpolluted areas. The lowest value was determined for and unpolluted site 2 in 2013 (2.99) and the highest for the polluted site 2 (4.38) in 2013 as well. The phylum-level Shannon diversity indices were not significantly different among samples, with 1.51 being the average value.

4.2 Bacterial expression system construction

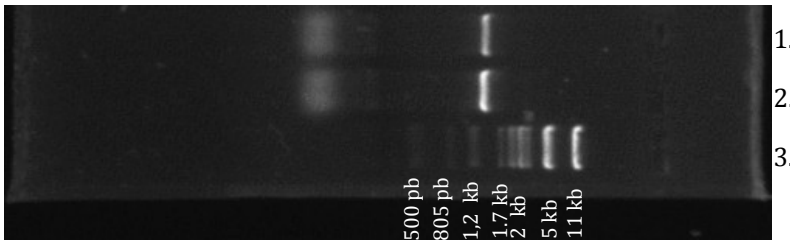
The aim of this part of the project was to identify potentially novel alpha-amino acid ester hydrolases among cultivable bacteria from the polluted soil and construct a bacterial expression system with the identified novel *aeh* gene.

First, the functional screening among cultivable bacteria was carried out using the chromogenic activity assays as described in Chapter 3.3.1. In total, 5 organisms with the desired phenotype were found, two of which, *Pantoea agglomerans* and *Bacillus cereus* (16S rDNA gene sequencing was used for bacterial identification), have not been previously described as AEH producers.

Molecular cloning of *aeh* gene

Chromosomal DNAs of both strains were isolated and the COnsensus- DEgenerate Hybrid Oligonucleotide Primer strategy using the previously described primers: AEH_Fwd190 and AEH_Rev670 was applied. No specific PCR product for the genomic DNA purified from *Pantoea agglomerans* was obtained. However, for *Bacillus cereus* the band of the desired size could be observed on agarose gel (see Figure 14). This 1450-bp fragment was cut out and purified using the High Pure PCR Product Purification Kit (see Chapter 3.3.5).

Figure 14: PCR results with degenerate primers and chromosomal DNA from *B. cereus*.
The bands in lanes 1 and 2 correspond to the product of duplicate PCR reactions and match the expected length of approx. 1.5 kb.



1. *aeh* from *B. cereus* gene fragment
2. *aeh* from *B. cereus* gene fragment
3. lambda DNA *Pst*I ladder

Subsequently, pK19 vector plasmid was opened to produce blunt ends with the use of *Sma*I restriction enzyme, the purified fragment of *aeh* gene was ligated into this vector and the ligation mixture was used to transform TOP 10 host, as described in Chapters 3.3.6, 3.3.7 and 3.3.8. The cells were plated onto selective medium supplied with kanamycin, x-gal and IPTG, as the vector allows testing the correct insertion of the fragment DNA by blue-white screening. Single white and light-blue colonies were transferred from the plates into 2 ml of LB medium with kanamycin and grown over night at 37 °C. A small scale plasmid purification was performed for individual samples and recombinant plasmids of matching size (validated on agarose gel with the original pK19 plasmid as a standard; figure not shown) were sequenced with the M13/pUC standard forward sequencing primer.

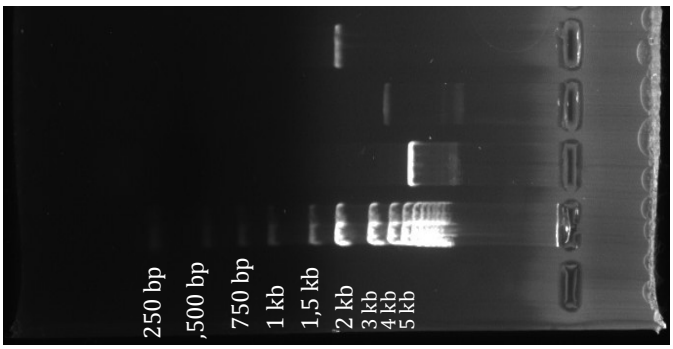
The acquired DNA sequence was as follows:

GGATTGCGGTCGTAGAGAGGGAACAGGCTGGACTGCACCTGCACCATCACCCGGTGGCCGCGCTCGAAGGTGTGGTTG
GCGGTGGGCAGGCCGAAGCTGTAGGCCAGCGGCTGGTTCGGGGTCAGCGGCGCCGGATGCTCGAAGCTCTCGCGGTAG
CGGCCGCGGAAGGTCGCCAGCGACACCGGCAGCTCGTAGCCGCCCAGCTTCGGCGCGGAGGCCATCTGGTCCGGGTAC
ACGTCGATCAGCTTCACCACCCAATCGCTGTCGCTGCCGCTGGCGGAGGCCTGCAGGTGCACCTGCGGCGCGCCGGCGA
TGCGCAGCGGCGCCTGCAGCGGCTCGCTGACGAAGGTCAGCACGTCCGGGCGCCCGTCGACGAAGCGCTGGTCGTGCA
CCAGCCAGGTGGTCCACATGTCGCGGTCGCCGAACACCACCGGGCGCGGCACGAACGGCACCGGCTTGGCCGGGTCGG
ACACGTACTCGGTGTACTCGCCCTGCCCGGCCTTGGGCGCCTCGAACGAGACGCGGCCGCCGGCCTCCAGGTACAGCGG
CTTGCTCTGCGCCGGGCAGCCCTGCGTGCAGCTCAGCGGCCAGCGCTGCAGGCGGTCCCAGTGGTTGGCGCCGGTGTCG
TAGATCAGCACCGGCGGGGTGTCGGCCTTGGGCGCGCCATCGACCAGGTACTGGTCGAAGAACGGCTTGAGCACGTCG
CGGCGGAACTGCAGTGCGGCGTCGCCGTCGAACTGCAGCGCGCCCAGCTAGGCGCCCTCGTAGTTGACCTGGCTATGC
CGCCACGGCCCCATCACCAGGTAGTTCTTGTCGTTGCCCGGTGTCGCGCGGCTCCATCGCCTCGTAGCTGTGGATCGCG
CCCCACATGTCTTCCCGGTCCCACAGCCCCTGCAGCCACATCGTCGGCACCTTTCAGCGGGGTGCGGCGCCATCACCTT
GTCCAGCGCTGCTCTGCAGACGCATCGTAGCCGGGTGCTCGGTGAGCTGTGCACACGCCACTGCTCAGC

The sequence was compared to those deposited in the GenBank database, and it was found that the gene fragment has a high level of nucleotide alignment (98%) with *aeh* gene from *Acidovorax avenae* (synonym for *Xanthomonas rubrilineans*) strain SIIA (Qu et al., 2012; Pan et al., 2013). Therefore, the original strategy to acquire the flanking regions of *aeh* by inverse PCR was put aside and a pair of specific primers was designed according to the known *aeh* sequence from *X. rubrilineans:* Fwd_AEH_Nde_ATG – forward primer, beginning with the start codon and containing *Nde*I restriction site; and Rev_AEH_Xho_Stop – reverse primer, ending with the termination codon and containing *Xho*I restriction site (see Table 1).

A new PCR amplification of the gene was performed on the genomic DNA of *B. cereus* using these specific primers and specific PCR products were extracted from agarose gel. The purified complete *aeh* gene (approx. 1917-bp long, with possibly mutated fragments at its start and end) was cut with both *Nde*I and *Xho*I in one reaction mixture carried out in Buffer O. The same digestion procedure was performed for pET26b vector plasmid. After that, both DNA samples were separated on agarose gel and examined for band size and quality (Figure 15).

Figure 15: Digested pET26b vector and digested insert (*Nde*I and *Xho*I restriction enzymes). Digested pET26b (lane 3) should correspond to approx. 5,2 kb, as 130-bp-long fragment between the two restriction sites was cut out.

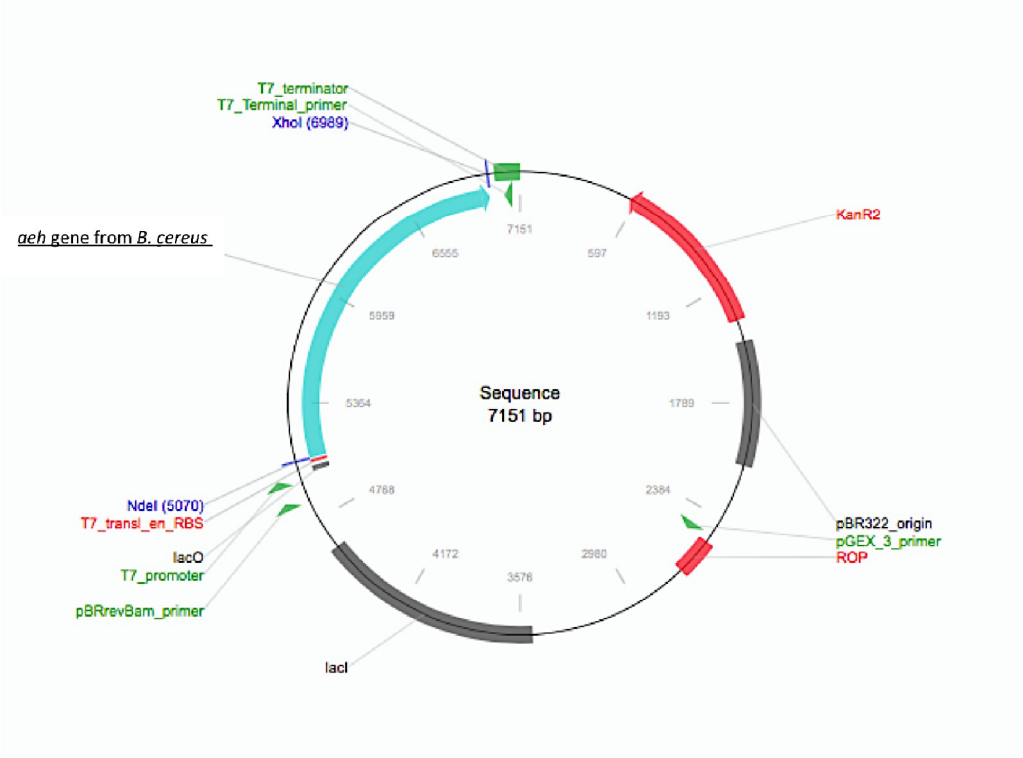


1. *aeh* gene digested with *Nde*I and *Xho*I
2. circular pET26b vector
3. pET26b vector linearized by *Nde*I and *Xho*I
4. 1 kb ladder

The corresponding products were excised and further purified using the High Pure PCR Product Purification Kit.

The *aeh* gene (about 2 kb) was then ligated overnight into pET26b through the complementary cohesive ends, ensuring the correct orientation towards T7 promoter (Figure 16). The expected recombinant plasmid was designated as pBCMN.

Figure 16: A circular map of the recombinant plasmid pBCMN with the relevant restriction sites. Plasmid map was created with Addgene online tool http://www.addgene.org/analyze-sequence/ (27.03.2015)



The ligation mixture was used to transform the host *E. coli* TOP 10 because the construct seemed rather large for direct transformation of *E. coli* BL21 (DE3). The competent *E. coli* TOP 10 cells were transformed by the heat shock with the complete 20-µl ligation mixture and grown on kanamycin plates overnight. Vector pET26b plasmid does not allow for blue-white screening, and the hydrolase phenotype could not be verified in TOP 10 cells, which do not carry T7 RNA polymerase in their chromosome. Therefore the acquisition of correct construct was validated by purification of the plasmids and comparing their size on agarose gel and subsequent Colony PCR for selected colonies, again followed by visualisation on agarose gel.

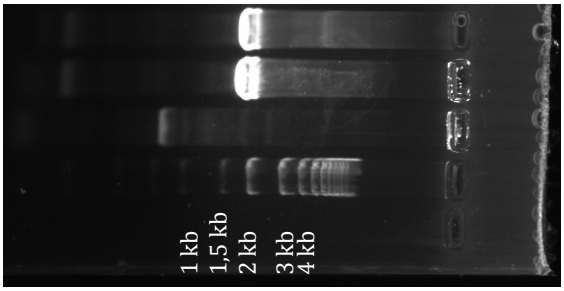
Among all tested constructs, plasmids number 16 and 20 and possibly 17 migrated more slowly in the gel than the vector control did, indicating that they contained an insert (Figure 17).

Figure 17: Small scale plasmid purification to verify the correct construct.
Plasmids number 16 and 20 and possibly 17 (underlined) migrated more slowly than pET26b vector, indicating that they contained an insert.

pET26b vector
sample 24
sample 23
sample 22
sample 21
sample 20
sample 19
sample 18
sample 17
sample 16
sample 15
sample 14
sample 13

The results of the colony PCR with specific primers Fwd_AEH_Nde_ATG and Rev_AEH_Xho_Stop suggested that colonies number 16 and 20 contain the correct construct (Figure 18).

Figure 18: The results of the colony PCR with specific primers.
The expected product length is approx. 2 kb. The size of samples number 20 and 16 (underlined) corresponds to the expected.

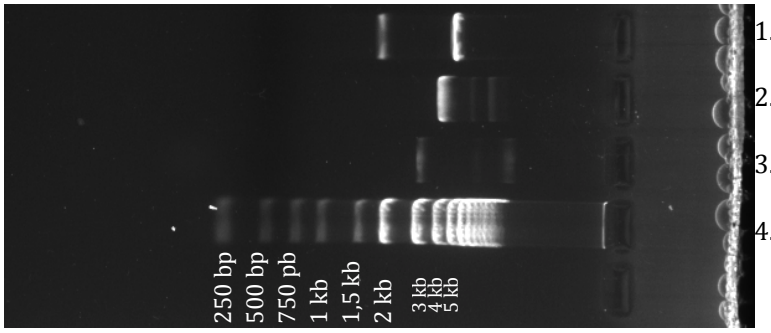


1. sample 20
2. sample 16
3. sample 17
4. 1 kb ladder

Both plasmids were sent to an external facility for sequencing with the standard pET primers, which revealed that a mutation occurred in the start codon region of plasmid number 16. As concerns the construct number 20, the sequencing confirmed that the *aeh* gene is complete (see Supplement 1 for the sequence) and correctly oriented. Therefore this construct represents the plasmid pBCMN and was selected for further cloning into BL21 (DE3) cells.

The plasmid number 20 was also validated by restriction: the insert was cut out with the use of *Nde*I and *Xho*I enzymes and consequently the rest of the vector was linearized. The agarose gel with restriction fragments is shown in Figure 19, lane 1.

Figure 19: Verification of plasmid 20 by digestion.
The insert (*aeh* gene) is expected to be cut out and form a 2 kb-long band on gel. The size of the linearized vector is presumed to be 5 kb. Both expected products can be observed in lane 1.



1. construct 20, digested with *Nde*I and *Xho*I
2. circular construct 20
3. pET26b vector
4. 1 kb ladder

As can be seen the bands in the lane 1 are about 2 kb and 5 kb in length, indicating that the vector contained an insert of a correct size cloned through the abovementioned restriction sites before it was linearized.

The strain *E.coli* TOP10 (pBCMN) was cultivated overnight in LB medium with kanamycin, the plasmid was purified using the High Pure Plasmid Isolation Kit and eluted to 30 µl of Elution buffer, of which 3 µl were used to transform the host *E. coli* BL21 (DE3). Among the grown up kanamycin-resistant colonies, 24 were picked from the plates for functional screening using the chromogenic activity assays on a culture plate. Four colonies with the highest activity (rich yellow colour, evaluated by eye) were selected for further kinetic assays, conducted as described in Chapter 3.3.13. The enzyme assay was also conducted for *B. cereus*, as a native AEH producer (highlighted in Table 3).

Table 3: The results of the enzyme assays for shaken flask cultivation of four clones of BL21(DE3)(pBCMN) and nature isolate *B. cereus*.

Sample	CDW (g/L)	VA (U/L)	SA (U/g_{cdw})
B. cereus	3,0	1,5	0,5
20/I	5,4	0	0
20/II	6,0	6314	1052
20/III	8,2	0	0
20/IX	5,7	1 110	195

Isolate number 20/II exhibited the highest activity in shaken flask kinetic assays, therefore it was selected for further characterisation and production study in a stirred bioreactor. The data from bioreactor are presented in Table 4.

Fed-batch culture of the strain BL21(DE3)(pBCMN) in a bioreactor

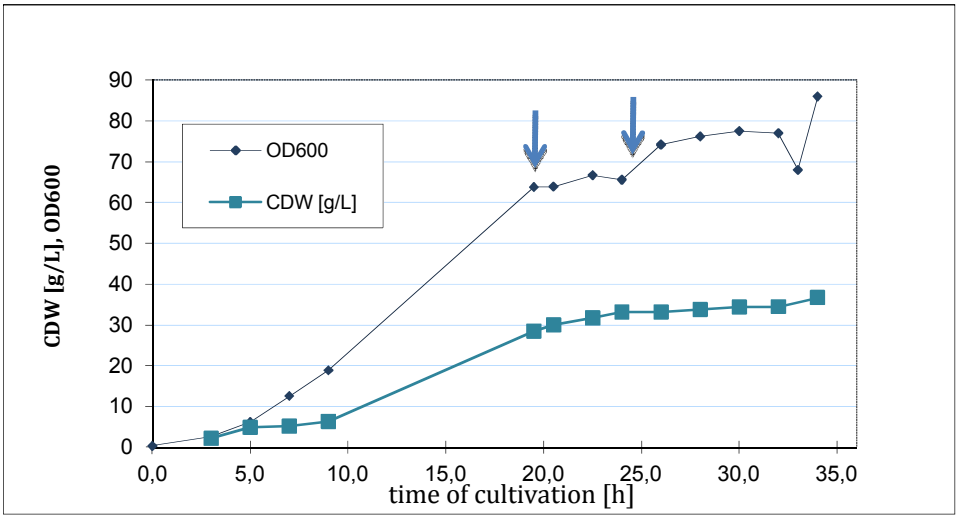
The culture was performed in a continuously stirred bioreactor with starting working volume of 6 L of medium MYEGly. Table 4 shows the time course of OD_{600}, CDW, VA and SA.

Table 4: The fed-batch cultivation of BL21(DE3)(pBCMN) in stirred bioreactor. Time-courses of biomass concentration (OD_{600}, CDW), residual concentration of inducer (lactose) and specific (SA) and volumetric (VA) activities of AEH.

	hour	OD_{600}	CDW (g/L)	lactose/100 in medium (g/L)	VA (U/L)	SA (U/g_{cdw})
	0.0	0.328				
	3.0	2.65	2.2			
	5.0	6.24	4.9			
	7.0	12.60	5.2			
	8.0					
	8.5					
	9.0	18.90	6.4			
1.induction	19.5	63.80	28.5	27.60	453	160
	20.5	63.90	30.0	135.20	548	180
	22.5	66.70	31.7	84.60	2668	840
	24.0	65.60	33.2	93.00	6063	1830
2.induction	25.0					
	26.0	74.20	33.2	106.40	12625	3800
	28.0	76.20	33.8	92.60	16350	4484
	30.0	77.50	34.4	110.60	25625	7449
	32.0	77.00	34.5	87.60	25313	7340
	34.0	86.00	36.7	141.40	27375	7460

The culturing was started as a batch culture. After about 10 hours of growth, feeding of glycerol was initiated. At OD_{600} of 40, synthesis of AEH was induced by a dose of lactose. To maintain inducer concentration above the value of 80 g/L, second dose of lactose was applied at 36th hour of growth. Final culture parameters were as follows: CDW 36,7 g/L, VA 27375 U/L and SA about 7460 U/L.

Figure 20: Fed-batch cultivation of E. coli BL21(DE3)(pBCMN) – culture growth. Arrows indicate time of dosing of inducer lactose.



5 Discussion

Microbial consortia of industrially polluted soils

Approximately 1% of the total soil bacterial population can be cultured by standard laboratory practices. DNA of residual 99% therefore represents a huge pool of novel, interesting structural genes. To acquire an access to these genes, an isolation of DNA from soil samples can be performed using various protocols, which can be classified as direct or indirect DNA extraction procedures. Direct DNA isolation is based on cell lysis within the sample matter and subsequent separation of DNA from the matter and cell debris (Ogram et al., 1987). The indirect approach, on the contrary, starts with the separation of cells from the environmental material prior to the lytic release of DNA (Holben et al., 1988). The indirect approach is acknowledged to yield DNA of higher molecular mass and greater purity than direct lysis procedures. However, in many cases the amounts of DNA recovered by this strategy are significantly lower in comparison to direct lysis protocols. As both methods have their advantages and drawbacks, the environmental DNA in this study was isolated for each sample by both methods, with the products subsequently pooled together for each of the samples.

Since the late 1930s, when the antibiotics were introduced, their usage has increased rapidly for human and animal treatment. Most antibiotics are considered "pseudo-persistent" (Li, 2014), which means they enter into the environment continuously and lead to permanent presence. Among them, beta-lactam antibiotics, widely used for human therapy, have been produced for many years nearby the sampling sites chosen for this study. The presence of antibiotics in the environment may represent a selection pressure on microbial consortia of polluted areas.

There is no consensus in scientific community for what constitutes a microbial species, that is why the operational taxonomic unit is a widely used construct of clustered sequence data that approximates "species" in subsequent analysis steps. In present thesis the sequences were clustered into OTUs, their identification was performed by BLASTn against local database derived from Ribosomal Database Project and each OTU was assigned to the taxonomic level of genus.

Even though all acquired data went through stringent quality controls, there are still some potential biases introduced by the choice of hypervariable region to amplify, as it has been shown that different regions of 16S rRNA gene reveal different diversity, and therefore a certain region may serve well for profiling a certain spectrum of bacteria but not all; other limitations of the method were previously discussed in Literature Review section. Some of the potential biases, including PCR mutation and chimera, cause an overestimation of richness, while e.g. primer mismatching leads to an underestimation. Therefore, it would be inappropriate to directly compare the taxon richness from this study the results obtained using different protocols.

On the other hand, the Shannon diversity index is much less sensitive to sequencing and PCR errors as well as to sampling depths. It could be deduced from its formula that inputs at low

frequencies, either from undetermined rare species or from experimental errors, contribute little to the Shannon index, it is less sensitive to experimental conditions of PCR and the use of different sequencing platforms (e.g. pyrosequencing) (Huse et al., 2010). Shannon index reflects both the number of species present and the evenness of their distribution.

As can be derived from the Shannon diversity index formula (see Chapter 3.2.6), high values of H' would be representative of more diverse communities. A community with only one species would have an H' value of 0 because P_i would equal 1 and be multiplied by ln P_i, which would equal zero. If the species are evenly distributed then the H' value would be high. Therefore the H' value allows us to know not only the number of species, but how the abundance of the species is distributed among all the species in the community. Significant differences in Shannon diversity index values could be observed between groups, and contrary to the initial assumption, the diversity was higher in polluted areas than in unpolluted.

From the biological point of view, the index can be converted into the effective number of species (ENS), which allows comparing the biodiversity with other communities. A community with Shannon index of H' has an equivalent diversity as a community with equally-common species of exp(H'), the ENS. Thus, the Shannon index of 2.99 (the lowest determined for the samples in this thesis, unpolluted site 2 in 2013) corresponds to ENS ≈ 20, meaning that this community has an equivalent diversity as a community with 20 equally-common species. The community with the highest measured Shannon index in the present thesis, coming from the polluted site 2 sampled in 2013, has an equivalent diversity as a community with 80 equally-common species. Typical values of Shannon index in most ecological studies are generally between 1.5 and 3.5, and the index is rarely greater than 4 (Magurran, 2004).

At genus level, the differences between the communities from studied areas were detected by Venn diagram (data not shown). Among the 100 most abundant genera (genera with the total relative abundance over 0.1%) more than 95% of taxa belong to the shared genera. Differences thus can be found only among the less abundant species. A genus-level comparison to other studies is hardly possible; on the other hand, the group of dominant phyla (Proteobacteria, Acidobacteria, Verrucimocrobia, Actinobacteria and Bacterioides) is typical for the soil samples collected globally, with their relative abundances, though, often varying 38-fold among those samples (Ramirez et al., 2014).

Principal component analysis was applied to identify the community structure differences in different locations. PCA tools are particularly effective when the variable of interest is associated with major changes in community structure, but are less effective at detecting subtle variations (Preheim et al., 2013). The resolution of PCA between phylum, genus and OTU levels can be different, e.g. see (Hong et al., 2015). For the purposes of this study the analysis was performed at

the genus level to achieve better discrimination. The first and the second components explained 79.7% and 9.7% of variability, respectively, accounting for over 89% of the variation in total.

The analysis demonstrated that although the communities dwelling in polluted and unpolluted areas can be separated, there is no clear boundary, and the unpolluted samples collected in 2014 are approaching the polluted ones. Notably well distinguished distance between all polluted samples and the unpolluted ones from the year 2013 could be observed. The group of unpolluted samples from the year 2014 was, except for one replicate of the unpolluted 2 site, also gathered together and plotted closer to the polluted samples, which leads to the hypothesis that external influences led to the spread of the contaminants.

The PCA at genus level also revealed that the significant gap between unpolluted samples from 2013 and the polluted areas was in accordance with the predominance of *Aeromicrobium* and *Salmonella* in the sample. Other samples are mainly differentiated from those by the presence and abundance of the *Halochromatium*, *Acidobacterium*, *Desulfomonile* (mainly unpolluted samples from 2014 and partly polluted samples site 2 from 2014) and *Gemmetimonas* and *Rubrivivax* (polluted samples from both 2013 and 2014). Many research studies have confirmed that this phenomenon had close relationship with the acidic nature of the polluted soil (Hong et al., 2015). However, in the present study, the Non-metric Multidimensional Scaling analysis, which was performed to identify the relationships between the most abundant genera and pH of the soil, determined that pH did not significantly impact the microbial community across all samples at genus level (data not shown). The assumption concerning this phenomenon turns to other characteristics of the soil, namely to possibly increased concentration of the salts in polluted samples, which could be caused by the industrial production. This hypothesis is based on the fact that prevalent bacteria in polluted samples have generally higher tolerance of the salts in comparison to the prevalent genera in unpolluted samples collected in 2013. The high salinity of the soil on microbial consortia is known to have high ecological impact (Bongoua Devisme, 2014). Additional measurement of the soil salt concentrations is therefore necessary.

A year-on-year statistical comparison by MANOVA between samples collected in 2013 and 2014 showed that significant difference could be seen among the unpolluted samples; their composition appeared to be on the border between all polluted samples and the unpolluted ones from 2013. This indicates that external influences led to the spread of the contaminants. It should be noted that the sampling in 2013 was performed in the non-manufacturing period (beta-lactam antibiotics were not produced over the year before the sampling); the sampling in 2014 was performed after 6 months of Penicillin G production. It can be therefore assumed that the potential impact of this industrial production is not exclusively of tong-term character, but even shorter manufacturing periods can remarkably impact microbiome.

In conclusion, this work represents a pilot study of the bacterial community variations on genus level under industrial-caused pollution. It demonstrates significant differences between polluted and unpolluted samples, and therefore justifies the need for additional experiments and statistical analyses to fully understand the impact of the production of beta-lactam antibiotics on the composition and ecology of the soil bacterial communities of adjacent area.

Analysis of bacterial community is typically conducted in one of the following three ways:

- Culturing bacteria from an environmental sample. Thus, before the application of next-generation sequencing, the genes of interest were typically isolated by cloning from cultured bacteria or by PCR amplification. However culture-based analysis excludes most of the microbial diversity, and PCR detection depends on the knowledge of gene sequence and doesn't allow for the discovery of novel genes.
- To access a broader suite of genes, sequence-based metagenomics can be helpful. It involves extracting and random sequencing of DNA directly from the environment without the need for culturing. The environmental DNA is sequenced and explored as to occurrence of gene sequences related to known genes of interest. Several novel enzymes, for instance belonging to beta-lactamases (Lee and Lee, 2011), chitinases (LeCleir et al., 2004) and epoxide hydrolases (Kotik, 2009) were retrieved from environmental samples using this sequence homology strategy.
- As a third approach, functional metagenomics can be applied, a technique that enables discovery of genes from organisms that may or may not be readily cultivable. Selections for functional enzymes can interrogate all expressed proteins in the library and not just those with sequences homologous to known proteins, thus revealing the genes that otherwise would be missed by sequence analysis alone. Thus, using this approach, a research group from the University of Wisconsin have found the genes encoding two putative novel functional enzymes. One of them was a bifunctional protein conferring resistance to ceftazidime (Donato et al., 2010).

The present thesis is in fact a part of a more extensive research, exploring the bacterial community of pharmaceutically polluted soil as such, investigating the resistome and looking for industrially relevant biocatalysts by all three mentioned approaches.

Thus, among cultivable bacteria from the environmental sample, 5 organisms with AEH phenotype were found. Two of them, taxonomically classified as *Pantoea agglomerans* and *Bacillus cereus,* have not been previously described as AEH producers. The respective gene identification by the CODEHOP primer design strategy was unsuccessful for *P. agglomerans,* suggesting that the DNA sequence of this novel enzyme is not homologous to known AEH proteins and the gene

should be "mined" using a different strategy, e.g. creating a small genomic library with subsequent functional screening.

Bacterial expression system construction

The fragment of *aeh* gene form *B. cereus* was successfully obtained with the use of degenerate primers, ligated to pK19 vector plasmid, and the host *E.coli* TOP10 cells were transformed with the ligation mixture. The fragment DNA sequence was compared to those deposited in the GenBank database, and it was found that it embodies a high level of nucleotide alignment (98%) with *aeh* gene from *Acidovorax avenae* (synonym for *Xanthomonas rubrilineans*). Therefore, the original strategy to acquire the flanking regions of *aeh* by inverse PCR was given up and the complete gene was retrieved by PCA with specific primers designed according to the known *aeh* sequence from *X. rubrilineans.*

Cohesive end ligation of the complete gene to pET26b expression vector was performed. After transformation of *E.coli* TOP 10 with the above-mentioned ligation mixture, the construct was verified as to correct gene insertion. Then *E.coli* BL21 (DE3), the system that allows to separate in time the growth phase and the expression of the desired product due to the control of T7 promoter, was used as a novel host and the validated recombinant plasmid was designated as pBCMN.

The activity assays performed for flask and fed-batch cultivations of *E.coli* BL21 (DE3) (pBCMN) revealed that specific hydrolytic activity for fed-batch culture was seven times higher than that for the most active shaken flask culture (7460 U/g_{cdw} compared to 1052 U/g_{cdw}). Fed-batch cultivation can be defined as an operational technique in biotechnological processes where one or more nutrients are supplied to the bioreactor during cultivation. Consequently, the volume of the culture increases and the biomass can achieve much higher concentrations. It is assumed, that fed-batch cultivation is superior to conventional batch culture, because controlling concentrations of a nutrient affect the yield and productivity of the desired metabolite (Yang, 1992). The results of the present study are in compliance with this assumption.

Alpha-amino acid ester hydrolases in general possess high industrial potential for biocatalysis of beta-lactams. Many examples of SSBAs syntheses catalysed by AEHs can be found in literature (see Table 5). Thus, AEH from *X. rubrillineans* was expressed heterologously in *E. coli* BL21 (DE3) and utilized for cefatrizine synthesis with 64.3% transformation yield (Pan et al., 2013). Likewise, AEH from *X. citri* was used for amoxicillin synthesis with the yield around 60% (Kato et al., 1980). The enzyme from *B. cereus* presented in this research will be further characterized and compared with the AEHs described in other studies. Eventually, the enzyme characteristics can be improved by site-directed mutagenesis, since there are examples of such improvement of AEH as to increase of the ratio of synthesis to hydrolysis, achieving higher maximum product accumulation (Barends et al., 2006) and improvement of enzyme thermostability (Blum et al., 2012).

Table 5: Semi synthetic beta-lactam antibiotics synthesized with AEH, adapted from (Marešová et al., 2014)

Antibiotic	Beta-lactam nucleus	Acyl donor	Used by
Ampicillin	6-APA	PGME	(Fernandez-Lafuente et al., 2001); (Blum and Bommarius, 2010)
Amoxicillin	6-APA	Hydroxy-PGME	(Kato et al., 1980)
Cefaclor	7-ACCA (7-Aminochloro-cephemcarboxylic Acid)	PGME	(Wang et al., 2012)
Cafadroxil	7-ADCA	Hydroxy-PGME	(Ye et al., 2012)
Cephalexin	7-ADCA	PGME	(Takahashi et al., 1974); (Polderman-Tijmes et al., 2002); (Blinkovsky and Markaryan, 1993); (Barends et al., 2006); (Blum and Bommarius, 2010)
Cefprozil	7-APRA (7-aminopropenyl-cephemcarboxylic acid)	Hydroxy-PGME	(Ye et al., 2012)
Cephaloglycin	7-ACA	PGME	(Takahashi et al., 1972)
Cefatrizine	7-ATTC (7-amino-3-methyl-cephalosporanic acid)	Hydroxy-PGME	(Pan et al., 2013); (Ye et al., 2012)

In the present study, after a complete DNA sequence of *aeh* from *B. cereus* was retrieved, it was found out that the level of identity with *aeh* from *X. rubrillineans,* is very high (around 99%). Also, all key catalytic residues were conserved including the catalytic triad, carboxylate cluster and oxyanion hole. These results lead to a hypothesis that the studied *aeh* gene could have been acquired by *B. cereus* by horizontal transfer, which has not been previously described for alpha-amino acid ester hydrolases. Another project performed in the laboratory is focused on this subject.

In general, bacteria can gain new properties by mutating existing genes (vertical evolution) (Martinez and Baquero, 2000), or by acquiring new genes from other strains or species (horizontal gene transfer) (Hegstad et al., 2010). The sharing of genes between bacteria by horizontal gene transfer occurs by many different mechanisms. This transfer can be mediated by mobile genetic elements, including phages, plasmids and transposons mediate this transfer. Also, it was concluded by a research team from China studying the genome of a non-pathogenic phenol-degrading bacterium, that many genes associated with environmental adaptation were acquired by horizontal gene transfer, including an 8 kb phenol degradation gene cluster (Zhan et al., 2012). It was further discovered, that in some circumstances the presence of low levels of the antibiotic in the environment is the key signal that promotes gene transfer, probably to ensure this way, that the whole microbial community is protected from the antibiotic. (Jeters et al., 2009).

At present, the biological function of AEH remains unclear. The *aeh* gene presented in this study comes from a cultivable organism, which is advantageous for the examination of its biological function. It is known from the investigation of genetic organization of the DNA region in which the *aeh* gene is situated conducted for *A. turbidans*, that it appears be located in an area of genes involved in amino compounds metabolism (Polderman-Tijmes et al., 2002).

The physiological role of another enzyme utilized for SSBA synthesis, penicillin G acylase from *E. coli,* has been suggested to be involved in the degradation of phenylacetylated compounds for the generation of phenylacetic acid as a carbon source (Valle et al., 1991). Intriguingly, PGA from *Kluyvera citrophila* was found to have an ability to cleave bacterial quorum sensing signal molecules, acyl homoserine lactones (the largest and the most well characterized class of signal molecules) (Mukherji et al., 2014). And vice versa, acylhomoserine lactone acylase from *Streptomyces sp*. was found to be capable of degrading penicillin G by deacylation (Park et al., 2005). The biological function of PGA and AEH might be related to a possible signaling role of antibiotics at subinhibitory concentrations (Davies et al., 2006). Quorum sensing has been shown to regulate a variety of important cellular functions like mating and virulence against the host; therefore further research on the biological function of the enzymes presented in this study would be beneficial.

Conclusion

Present study is the first metagenomic characterisation of the microbiome of soils exposed to the long-term manufacturing of beta-lactam antibiotics; the study also shows that the metagenome of the polluted environments can serve as a great source of novel genes of interest.

Analysis of V4 region of the bacterial communities from the studied samples revealed high complexity of soil microbial environment and confirmed that anthropogenic activity (represented by production of beta-lactam antibiotics) notably influences the variability and abundance of the species, yet without reducing the microbial diversity.

Ilumina sequencing highlighted the predominance of the *Blastocatella* and *Gemmatimonas* and some other genera in the polluted soil samples in comparison to the unpolluted samples. This indicates that the above-mentioned bacteria could have possibly evolved the mechanism enabling them to overcome the selection pressure of the pollutant. However, many of the taxa, such as *Gemmatimonas* or *Rurivivax* remain poorly characterized at physiological and functional levels, justifying future research to obtain a comprehensive view of their role. The results also show that minor changes in soil pH do not necessarily trigger a change in the microbial community composition, proving the hypothesis that described differences between the polluted and unpolluted samples are caused by anthropogenic activity.

Soil samples of the polluted area were also subjected to a functional analysis in the search of novel alpha-amino acid ester hydrolase encoding genes. This enzyme has been proven to be able to hydrolyse semi-synthetic beta-lactam antibiotics and, even though its biological function remains unknown, there was an assumption of increased abundance and diversity of this enzyme in studied samples. Alpha-amino acid ester hydrolases, in general, possess high industrial potential for biocatalysis of beta-lactam antibiotics and therefore discovering novel ones is of particular interest.

Among cultivable bacteria from polluted soil, five organisms with alpha-amino acid ester hydrolase phenotype were identified. Two of them, *Pantoea agglomerans* and *Bacillus cereus*, have not been previously described as AEH producers. The gene from *P. agglomerans* could not be retrieved with the use of degenerate primers, suggesting that the DNA sequence of this novel enzyme is not homologous to known AEH proteins. The full-length *aeh* gene from *B. cereus* was successfully obtained. The level of identity with already described *aeh* gene from *X. rubrillineans,* was very high which lead to a hypothesis that presence of the obtained gene in *B. cereus* is a result of the horizontal transfer, which has not been to date described for this particular gene. By molecular cloning and preparation of the recombinant production system (pET system-based, *E.coli* BL21 (DE3)) the efficient production platform for this enzyme was prepared. Data show that the obtained gene has the potential for future applications in biotechnologies. The present study illustrates both, how the soil microbial communities are influenced by industrial production and their ability to serve as a source of novel, potentially biotechnologically applicable, products.

References

Al-Ahmad, A., Daschner, F. D. and Kümmerer, K. (1999). Biodegradability of Cefotiam, Ciprofloxacin, Meropenem, Penicillin G, and Sulfamethoxazole and Inhibition of Waste Water Bacteria. Archives of Environmental Contamination and Toxicology 37(2): 158-163.

Alexander, M. (1978). Introduction to Soil Microbiology. Soil Science 125(5).

Altermann, E. (2012). Tracing Lifestyle Adaptation in Prokaryotic Genomes. Frontiers in Microbiology 3: 48.

Arnold, B. H., Filders, R. A., Gilbert, D. A. (1968). Cephalosporin derivatives. US patent 3821209.

Asai, T., Zaporojets, D., Squires, C. and Squires, C. L. (1999). An Escherichia coli strain with all chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria. Proceedings of the National Academy of Sciences of the United States of America 96(5): 1971-1976.

Baldwin, J. E. and Abraham, E. (1988). The biosynthesis of penicillins and cephalosporins. Nat Prod Rep 5(2): 129-145.

Barber, M., Giesecke, U., Reichert, A. and Minas, W. (2004). Industrial Enzymatic Production of Cephalosporin-Based Beta-Lactams. Molecular Biotechnolgy of Fungal beta-Lactam Antibiotics and Related Peptide Synthetases. A. A. Brakhage, Springer Berlin Heidelberg. 88**:** 179-215.

Barends, T. R., Polderman-Tijmes, J. J., Jekel, P. A., Hensgens, C. M., de Vries, E. J., Janssen, D. B. and Dijkstra, B. W. (2003). The sequence and crystal structure of the alpha-amino acid ester hydrolase from Xanthomonas citri define a new family of beta-lactam antibiotic acylases. J Biol Chem 278(25): 23076-23084.

Barends, T. R. M., Polderman-Tijmes, J. J., Jekel, P. A., Williams, C., Wybenga, G., Janssen, D. B. and Dijkstra, B. W. (2006). Acetobacter turbidans α-Amino Acid Ester Hydrolase: HOW A SINGLE MUTATION IMPROVES AN ANTIBIOTIC-PRODUCING ENZYME. Journal of Biological Chemistry 281(9): 5804-5810.

Bates, S. T., Berg-Lyons, D., Caporaso, J. G., Walters, W. A., Knight, R. and Fierer, N. (2011). Examining the global distribution of dominant archaeal populations in soil. ISME J 5(5): 908-917.

Bimboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Research 7(6): 1513-1523.

Blinkovsky, A. M. and Markaryan, A. N. (1993). Synthesis of β-lactam antibiotics containing α-aminophenylacetyl group in the acyl moiety catalyzed by d-(—)-Phenylglycyl-β-lactamide amidohydrolase. Enzyme and Microbial Technology 15(11): 965-973.

Blum, J. K. and Bommarius, A. S. (2010). Amino ester hydrolase from Xanthomonas campestris pv. campestris, ATCC 33913 for enzymatic synthesis of ampicillin. Journal of molecular catalysis. B, Enzymatic 67(1-2): 21-28.

Blum, J. K., Deaguero, A. L., Perez, C. V. and Bommarius, A. S. (2010). Ampicillin Synthesis Using a Two-Enzyme Cascade with Both α-Amino Ester Hydrolase and Penicillin G Acylase. ChemCatChem 2(8): 987-991.

Blum, J. K., Ricketts, M. D. and Bommarius, A. S. (2012). Improved thermostability of AEH by combining B-FIT analysis and structure-guided consensus method. Journal of Biotechnology 160(3–4): 214-221.

Boccu, E., Ebert, C., Gardossi, L., Gianferrara, T., Zacchigna, M. and Linda, P. (1991). Enzymatic synthesis of ampicillin: a chemometric optimization. Farmaco 46(4): 565-577.

Bongoua Devisme, A. J. (2014). Structure and Activity of Fe(III)-Reducing Microorganism Occurring in Paddy Fields of Thailand. International Journal of Sciences 06(2014):6-17.

Borneman, J. and Triplett, E. W. (1997). Molecular microbial diversity in soils from eastern Amazonia: evidence for unusual microorganisms and microbial population shifts associated with deforestation. Applied and Environmental Microbiology 63(7): 2647-2653.

Bowman, J. P., Rea, S. M., McCammon, S. A. and McMeekin, T. A. (2000). Diversity and community structure within anoxic sediment from marine salinity meromictic lakes and a coastal meromictic marine basin, Vestfold Hills, Eastern Antarctica. Environmental Microbiology 2(2): 227-237.

Brady, S. F. (2007). Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nat. Protocols 2(5): 1297-1305.

Brakhage, A. A. (1998). Molecular regulation of beta-lactam biosynthesis in filamentous fungi. Microbiol Mol Biol Rev 62(3): 547-585.

Brotzu, G. (1948). Research on a new antibiotic. Publications of the Cagliari Institute of Hygiene: 5–19.

Bruggink, A., Roos, E. C. and de Vroom, E. (1998). Penicillin Acylase in the Industrial Production of beta-Lactam Antibiotics. Organic Process Research & Development 2(2): 128-133.

Bruggink, A. E. (2001). Synthesis of Beta-lactam Antibiotics, Kluwer Academic Pub, Dordrecht.

Cantwell, C., Beckmann, R., Whiteman, P., Queener, S. W. and Abraham, E. P. (1992). Isolation of Deacetoxycephalosporin C from Fermentation Broths of Penicillium chrysogenum Transformants: Construction of a New Fungal Biosynthetic Pathway. Proceedings of the Royal Society of London. Series B: Biological Sciences 248(1323): 283-289.

Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R. and Gilbert, J. A. (2012). The Western English Channel contains a persistent microbial seed bank. ISME J 6(6): 1089-1093.

Cardinali-Rezende, J., Debarry, R., Colturato, L. D. B., Carneiro, E., Chartone-Souza, E. and Nascimento, A. A. (2009). Molecular identification and dynamics of microbial communities in reactor treating organic household waste. Applied Microbiology and Biotechnology 84(4): 777-789.

Cole, J. R., Chai, B., Farris, R. J., Wang, Q., Kulam, S. A., McGarrell, D. M., Garrity, G. M. and Tiedje, J. M. (2005). The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Research 33(Database issue): D294-D296.

Conn, H. J., Harding, H. A., Kligler, I. J., Frost, W. D., Prucha, M. J. and Atkins, K. N. (1918). Methods of pure culture study preliminary report of the committee on the chart for identification of bacterial species. Journal of Bacteriology 3(2): 115-128.

Crawford, L., Stepan, A. M., McAda, P. C., Rambosek, J. A., Conder, M. J., Vinci, V. A. and Reeves, C. D. (1995). Production of cephalosporin intermediates by feeding adipic acid to recombinant Penicillium chrysogenum strains expressing ring expansion activity. Biotechnology (N Y) 13(1): 58-62.

D'Costa, V. M., Griffiths, E. and Wright, G. D. (2007). Expanding the soil antibiotic resistome: exploring environmental diversity. Current Opinion in Microbiology 10(5): 481-489.

Dane, E. and Dockner, T. (1964). Synthesis of 6-[(D-α-Amino-α-phenylacetyl)amino]-penicillanic Acid Using β-Dicarbonyl Compounds as Amino-Protecting Groups. Angewandte Chemie International Edition in English 3(6): 439-439.

Daniel, R. (2004). The soil metagenome – a rich resource for the discovery of novel natural products. Current Opinion in Biotechnology 15(3): 199-204.

Davies, J., Spiegelman, G. B. and Yim, G. (2006). The world of subinhibitory antibiotic concentrations. Current Opinion in Microbiology 9(5): 445-453.

de la Torre, A., Iglesias, I., Carballo, M., Ramírez, P. and Muñoz, M. J. (2012). An approach for mapping the vulnerability of European Union soils to antibiotic contamination. Science of The Total Environment 414(0): 672-679.

DeAngelis, K. M., Pold, G., Topçuoğlu, B. D., van Diepen, L. T. A., Varney, R. M., Blanchard, J. L., Melillo, J. and Frey, S. D. (2015). Long-term forest soil warming alters microbial communities in temperate forest soils. Frontiers in Microbiology 6: 104.

Delmont, T. O., Prestat E Fau - Keegan, K. P., Keegan Kp Fau - Faubladier, M., Faubladier M Fau - Robe, P., Robe P Fau - Clark, I. M., Clark Im Fau - Pelletier, E., Pelletier E Fau - Hirsch, P. R., Hirsch Pr Fau - Meyer, F., Meyer F Fau - Gilbert, J. A., Gilbert Ja Fau - Le Paslier, D., Le Paslier D Fau - Simonet, P., Simonet P Fau - Vogel, T. M. and Vogel, T. M. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. ISME J.(1751-7370 (Electronic)).

DeLong, E. F. (2005). Microbial community genomics in the ocean. Nat Rev Micro 3(6): 459-469.

Demain, A. and Elander, R. (1999). The beta-lactam antibiotics: past, present, and future. Antonie Van Leeuwenhoek 75(1-2): 5-19.

Demain, A. L. (1983). Biosynthesis of beta-Lactam Antibiotics. Antibiotics. A. Demain and N. Solomon, Springer Berlin Heidelberg. 67 / 1**:** 189-228.

Demain, A. L., Cho, H., Piret, J. M., Adrio, J. L., Fernandez, M.-J. E., Bàez-Vàsquez, M. A. and Hintermann, G. (2002). Penicillin conversion. US patent 6383773 B2.

Denome, S. A., Elf, P. K., Henderson, T. A., Nelson, D. E. and Young, K. D. (1999). Escherichia coli mutants lacking all possible combinations of eight penicillin binding proteins: viability, characteristics, and implications for peptidoglycan synthesis. J Bacteriol 181(13): 3981-3993.

Donato, J. J., Moe, L. A., Converse, B. J., Smart, K. D., Berklein, F. C., McManus, P. S. and Handelsman, J. (2010). Metagenomic Analysis of Apple Orchard Soil Reveals Antibiotic Resistance Genes Encoding Predicted Bifunctional Proteins. Applied and Environmental Microbiology 76(13): 4396-4401.

Duggleby, H. J., Tolley, S. P., Hill, C. P., Dodson, E. J., Dodson, G. and Moody, P. C. E. (1995). Penicillin acylase has a single-amino-acid catalytic centre. Nature 373(6511): 264-268.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Meth 10(10): 996-998.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27(16): 2194-2200.

Edmonds, J., Weston, N., Joye, S., Mou, X. and Moran, M. (2009). Microbial Community Response to Seawater Amendment in Low-Salinity Tidal Sediments. Microbial Ecology 58(3): 558-568.

Elander, R. P. (2003). Industrial production of beta-lactam antibiotics. Appl Microbiol Biotechnol 61(5-6): 385-392.

Faarup, P. (1977). Method of preparing a sparingly soluble complex of cephalexin. US patent 4003896 A.

Fernandez-Lafuente, R., Hernández-Justiz, O., Mateo, C., Terreni, M., Alonso, J., Garcia-López, J. L., Moreno, M. A. and Guisan, J. M. (2001). Stabilization of a tetrameric enzyme (α-amino acid ester hydrolase from Acetobacter turbidans) enables a very improved performance of ampicillin synthesis. Journal of Molecular Catalysis B: Enzymatic 11(4–6): 633-638.

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M., Edwards, R. A., Felts, B., Rayhawk, S., Knight, R., Rohwer, F. and Jackson, R. B. (2007). Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil. Applied and Environmental Microbiology 73(21): 7059-7066.

Fierer, N. and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. Proceedings of the National Academy of Sciences of the United States of America 103(3): 626-631.

Fierer, N., Leff, J. W., Adams, B. J., Nielsen, U. N., Bates, S. T., Lauber, C. L., Owens, S., Gilbert, J. A., Wall, D. H. and Caporaso, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proceedings of the National Academy of Sciences 109(52): 21390-21395.

Fleming, A. (1929). On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. Bull World Health Organ 79(8): 780-790.

Florey, H. W., Chain, E.B., Heatley, N.G., Jennings, M.A., Sanders, A.G., Abraham, E.P., Florey, M.E. (1949). Oxford, Oxford University Press.

Gans, J., Wolinsky, M. and Dunbar, J. (2005). Computational Improvements Reveal Great Bacterial Diversity and High Metal Toxicity in Soil. Science 309(5739): 1387-1390.

Gilbertson, T. J., Hornish, R. E., Jaglan, P. S., Koshy, K. T., Nappier, J. L., Stahl, G. L., Cazers, A. R., Nappier, J. M. and Kubicek, M. F. (1990). Environmental fate of ceftiofur sodium, a cephalosporin antibiotic. Role of animal excreta in its decomposition. Journal of Agricultural and Food Chemistry 38(3): 890-894.

Grulich, M., Štěpánek, V. and Kyslík, P. (2013). Perspectives and industrial potential of PGA selectivity and promiscuity. Biotechnology Advances 31(8): 1458-1472.

Gullberg, E., Cao, S., Berg, O. G., Ilbäck, C., Sandegren, L., Hughes, D. and Andersson, D. I. (2011). Selection of Resistant Bacteria at Very Low Antibiotic Concentrations. PLoS Pathogens 7(7): e1002158.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemistry & Biology 5(10): R245-R249.

Harris, D. M., Westerlaken, I., Schipper, D., van der Krogt, Z. A., Gombert, A. K., Sutherland, J., Raamsdonk, L. M., van den Berg, M. A., Bovenberg, R. A., Pronk, J. T. and Daran, J. M. (2009). Engineering of Penicillium chrysogenum for fermentative production of a novel carbamoylated cephem antibiotic precursor. Metab Eng 11(2): 125-137.

Hayes, J. D. and Wolf, C. R. (1990). Molecular mechanisms of drug resistance. Biochem J 272(2): 281-295.

Hegde, M. M., Thadani, S. B., Singh, U. and Naik, S. R. (1997). Isolation and purification of penicillin G acylase obtained from Escherichia coli (NCIM-2400) and immobilisation on Eupergit C for the production of 6 amino penicillanic acid. Hindustan Antibiot Bull 39(1-4): 1-10.

Hegstad, K., Mikalsen, T., Coque, T. M., Werner, G. and Sundsfjord, A. (2010). Mobile genetic elements and their contribution to the emergence of antimicrobial resistant Enterococcus faecalis and Enterococcus faecium. Clinical Microbiology and Infection 16(6): 541-554.

Henniger, P. W., van der Drift, J. K., VAN Veen, G. J. (1978). 6-(D-α-Amino-p-hydroxyphenylacetamido)-penicillanic acid preparation. US patent 4128547.

Holben, W. E., Jansson, J. K., Chelm, B. K. and Tiedje, J. M. (1988). DNA Probe Method for the Detection of Specific Microorganisms in the Soil Bacterial Community. Applied and Environmental Microbiology 54(3): 703-711.

Hong, C., Si, Y., Xing, Y. and Li, Y. (2015). Illumina MiSeq sequencing investigation on the contrasting soil bacterial community structures in different iron mining areas. Environmental Science and Pollution Research: 1-12.

Huse, S. M., Welch, D. M., Morrison, H. G. and Sogin, M. L. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environmental Microbiology 12(7): 1889-1898.

Janssen, P. H. (2006). Identifying the Dominant Soil Bacterial Taxa in Libraries of 16S rRNA and 16S rRNA Genes. Applied and Environmental Microbiology 72(3): 1719-1728.

Jeters, R. T., Wang, G.-R., Moon, K., Shoemaker, N. B. and Salyers, A. A. (2009). Tetracycline-Associated Transcriptional Regulation of Transfer Genes of the Bacteroides Conjugative Transposon CTnDOT. Journal of Bacteriology 191(20): 6374-6382.

Kasche, V. (1986). Mechanism and yields in enzyme catalysed equilibrium and kinetically controlled synthesis of Œ≤-lactam antibiotics, peptides and other condensation products. Enzyme and Microbial Technology 8(1): 4-16.

Kasche, V., Haufler, U., Markowsky, D., Melnyk, S., Zeich, A. and Galunsky, B. (1987). Penicillin Amidase from E. coli. Annals of the New York Academy of Sciences 501(1): 97-102.

Kasper, T. J., Melera, M., Gozel, P. and Brownlee, R. G. (1988). Separation and detection of DNA by capillary electrophoresis. Journal of Chromatography A 458(0): 303-312.

Kato, K., Kawahara, K., Takahashi, T. and Igarasi, S. (1980). Enzymatic Synthesis of Amoxicillin by the Cell-Bound α-Amino Acid Ester Hydrolase of Xanthomonas citri. Agricultural and Biological Chemistry 44(4): 821-825.

Katoh, K., Asimenos, G. and Toh, H. (2009). Multiple Alignment of DNA Sequences with MAFFT. Bioinformatics for DNA Sequence Analysis. D. Posada, Humana Press. 537**:** 39-64.

Kellenberger, E. (2001). Exploring the unknown: The silent revolution of microbiology. EMBO Reports 2(1): 5-7.

Klappenbach, J. A., Dunbar, J. M. and Schmidt, T. M. (2000). rRNA Operon Copy Number Reflects Ecological Strategies of Bacteria. Applied and Environmental Microbiology 66(4): 1328-1333.

Kohsaka, M. and Demain, A. L. (1976). Conversion of penicillin N to cephalosporin(s) by cell-free extracts of Cephalosporium acremonium. Biochemical and Biophysical Research Communications 70(2): 465-473.

Kormas, K. A., Tamaki, H., Hanada, S. and Kamagata, Y. (2009). Apparent richness and community composition of Bacteria and Archaea in geothermal springs. Aquatic Microbial Ecology 57(2): 113-122.

Kotik, M. (2009). Novel genes retrieved from environmental DNA by polymerase chain reaction: Current genome-walking techniques for future metagenome applications. Journal of Biotechnology 144(2): 75-82.

Kumar, R. R. and Prasad, S. (2011). Metabolic Engineering of Bacteria. Indian Journal of Microbiology 51(3): 403-409.

Kupka, J., Shen, Y.-Q., Wolfe, S. and Demain, A. L. (1983). Partial purification and properties of the α-ketoglutarate-linked ring-expansion enzyme of β-lactam biosynthesis of Cephalosporium acremonium. FEMS Microbiology Letters 16(1): 1-6.

Kurochkina, V. B., Sklyarenko, A. V., Berezina, O. V. and Yarotskii, S. V. (2013). Alpha-amino acid ester hydrolases: Properties and applications. Applied Biochemistry and Microbiology 49(8): 672-694.

Lauber, C. L., Hamady, M., Knight, R. and Fierer, N. (2009). Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. Applied and Environmental Microbiology 75(15): 5111-5120.

LeCleir, G. R., Buchan, A. and Hollibaugh, J. T. (2004). Chitinase Gene Sequences Retrieved from Diverse Aquatic Habitats Reveal Environment-Specific Distributions. Applied and Environmental Microbiology 70(12): 6977-6983.

Lee, S. H. and Lee, J.-H. (2011). Molecular Characterization of TEM-Type Beta-Lactamases Identified in Cold-Seep Sediments of Edison Seamount (South of Lihir Island, Papua New Guinea). Handbook of Molecular Microbial Ecology II, John Wiley & Sons, Inc.**:** 545-552.

Lee, S. Y., Kim, H. U., Park, J. H., Park, J. M. and Kim, T. Y. (2009). Metabolic engineering of microorganisms: general strategies and drug production. Drug Discov Today 14(1-2): 78-88.

Li, W. C. (2014). Occurrence, sources, and fate of pharmaceuticals in aquatic environment and soil. Environmental Pollution 187(0): 193-201.

Liras, P. and Martín, J. F. (2006). Gene clusters for b-lactam antibiotics and control of their expression: why have clusters evolved, and from where did they originate? International Microbiology 9: 09-19.

Liu, B., Faller, L. L., Klitgord, N., Mazumdar, V., Ghodsi, M., Sommer, D. D., Gibbons, T. R., Treangen, T. J., Chang, Y.-C., Li, S., Stine, O. C., Hasturk, H., Kasif, S., Segrè, D., Pop, M. and Amar, S. (2012). Deep Sequencing of the Oral Microbiome Reveals Signatures of Periodontal Disease. PLoS ONE 7(6): e37919.

Liu, Z., Lozupone, C., Hamady, M., Bushman, F. D. and Knight, R. (2007). Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Research 35(18): e120.

Magurran, A. E. (2004). Measuring Biological Diversity, Wiley.

Marengo, J. R., Kok, R. A., O'Brien, K., Velagaleti, R. R. and Stamm, J. M. (1997). Aerobic biodegradation of (14C)-sarafloxacin hydrochloride in soil. Environmental Toxicology and Chemistry 16(3): 462-471.

Marešová, H., Plačková, M., Grulich, M. and Kyslík, P. (2014). Current state and perspectives of penicillin G acylase-based biocatalyses. Applied Microbiology and Biotechnology 98(7): 2867-2879.

Marilley, L. and Aragno, M. (1999). Phylogenetic diversity of bacterial communities differing in degree of proximity of Lolium perenne and Trifolium repens roots. Applied Soil Ecology 13(2): 127-136.

Martin, C. H., Nielsen, D. R., Solomon, K. V. and Prather, K. L. (2009). Synthetic metabolism: engineering biology at the protein and pathway scales. Chem Biol 16(3): 277-286.

Martin, J. F. and Gutierrez, S. (1995). Genes for beta-lactam antibiotic biosynthesis. Antonie Van Leeuwenhoek 67(2): 181-200.

Martin, J. F. and Liras, P. (1989). Enzymes involved in penicillin, cephalosporin and cephamycin biosynthesis. Adv Biochem Eng Biotechnol 39: 153-187.

Martinez, J. L. and Baquero, F. (2000). Mutation Frequencies and Antibiotic Resistance. Antimicrobial Agents and Chemotherapy 44(7): 1771-1777.

Morin, R. B., Jackson, B. G., Flynn, E. H., Roeske, R. W. and Andrews, S. L. (1969). Chemistry of cephalosporin antibiotics. XIV. Reaction of cephalosporin C with nitrosyl Chloride. Journal of the American Chemical Society 91(6): 1396-1400.

Morin, R. B., Jackson, B. G., Mueller, R. A., Lavagnino, E. R., Scanlon, W. B. and Andrews, S. L. (1963). Chemistry of Cephalosporin Antibiotics. III. Chemical Correlation of Penicillin and Cephalosporin Antibiotics. Journal of the American Chemical Society 85(12): 1896-1897.

Mukherji, R., Varshney, N. K., Panigrahi, P., Suresh, C. G. and Prabhune, A. (2014). A new role for penicillin acylases: Degradation of acyl homoserine lactone quorum sensing signals by Kluyvera citrophila penicillin G acylase. Enzyme and Microbial Technology 56(0): 1-7.

Mukhopadhyay, R. (2009). DNA sequencers: the next generation. Analytical Chemistry 81(5): 1736-1740.

Newton, G. G. and Abraham, E. P. (1955). Cephalosporin C, a new antibiotic containing sulphur and D-alpha-aminoadipic acid. Nature 175(4456): 548.

Nieboer, M., de Vroom, E., Lugtenburg, J., Schipper, D., Vollebregt, A. W. H. and Bovenberg, R. A. L. (2003). Process for the fermentative production of deacylated cephalosporins. US patent 6518039.

Ogram, A., Sayler, G. S. and Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. Journal of Microbiological Methods 7(2–3): 57-66.

Ospina, S. S., Lopez-Munguia, A., Gonzalez, R. L. and Quintero, R. (1992). Characterization and use of a penicillin acylase biocatalyst. Journal of Chemical Technology & Biotechnology 53(2): 205-213.

Pace, N. R. (1997). A Molecular View of Microbial Diversity and the Biosphere. Science 276(5313): 734-740.

Pan, J., Wang, L., Li, D. and Ye, L. (2013). Synthesis of cefatrizine by recombinant alpha-amino acid ester hydrolase. Chinese journal of biotechnology(Apr;29(4)): 501-509.

Park, S.-Y., Kang, H.-O., Jang, H.-S., Lee, J.-K., Koo, B.-T. and Yum, D.-Y. (2005). Identification of Extracellular N-Acylhomoserine Lactone Acylase from a Streptomyces sp. and Its Application to Quorum Quenching. Applied and Environmental Microbiology 71(5): 2632-2641.

Parmar, A., Kumar, H., Marwaha, S. S. and Kennedy, J. F. (1998). Recent Trends in Enzymatic Conversion of Cephalosporin C to 7-Aminocephalosporanic Acid (7-ACA). Critical Reviews in Biotechnology 18(1): 1-12.

Polderman-Tijmes, J. J., Jekel, P. A., de Vries, E. J., van Merode, A. E. J., Floris, R., van der Laan, J.-M., Sonke, T. and Janssen, D. B. (2002). Cloning, Sequence Analysis, and Expression in Escherichia coli of the Gene Encoding an α-Amino Acid Ester Hydrolase from Acetobacter turbidans. Applied and Environmental Microbiology 68(1): 211-218.

Preheim, S. P., Perrotta, A. R., Friedman, J., Smilie, C., Brito, I., Smith, M. B. and Alm, E. (2013). Chapter Eighteen - Computational Methods for High-Throughput Comparative Analyses of Natural Microbial Communities. Methods in Enzymology. F. D. Edward, Academic Press. Volume 531**:** 353-370.

Pridmore, R. D. (1987). New and versatile cloning vectors with kanamycin-resistance marker. Gene 56(2–3): 309-312.

Qu, F., Yi B Fau - Ye, L. and Ye, L. (2012). Purification and characterization of alpha-amino acid ester hydrolase from Xanthomonas rubrillineans. Wei Sheng Wu Xue Bao(0001-6209 (Print)).

Ramirez, K. S., Leff, J. W., Barberán, A., Bates, S. T., Betley, J., Crowther, T. W., Kelly, E. F., Oldfield, E. E., Shaw, E. A., Steenbock, C., Bradford, M. A., Wall, D. H. and Fierer, N. (2014). Biogeographic patterns in below-ground diversity in New York City's Central Park are similar to those observed globally. Proceedings of the Royal Society of London B: Biological Sciences 281(1795).

Reddy, B. V. B., Kallifidas, D., Kim, J. H., Charlop-Powers, Z., Feng, Z. and Brady, S. F. (2012). Natural Product Biosynthetic Gene Diversity in Geographically Distinct Soil Microbiomes. Applied and Environmental Microbiology 78(10): 3744-3752.

Rose, T. M., Schultz, E. R., Henikoff, J. G., Pietrokovski, S., McCallum, C. M. and Henikoff, S. (1998). Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Research 26(7): 1628-1635.

Rousk, J., Baath, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R. and Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J 4(10): 1340-1351.

Rousk, J., Brookes Pc Fau - Baath, E. and Baath, E. (2011). Fungal and bacterial growth responses to N fertilization and pH in the 150-year 'Park Grass' UK grassland experiment. FEMS Microbiology Ecology 76 (1) (1574-6941 (Electronic)): 89-99.

Sanger, F. (1975). The Croonian Lecture, 1975: Nucleotide Sequences in DNA. Proceedings of the Royal Society of London B: Biological Sciences 191(1104): 317-333.

Schmieder, R. and Edwards, R. (2011). Insights into antibiotic resistance through metagenomic approaches. Future Microbiology 7(1): 73-89.

Schroen, C. G. P. H., Nierstrasz, V. A., Kroon, P. J., Bosma, R., Janssen, A. E. M., Beeftink, H. H. and Tramper, J. (1999). Thermodynamically controlled synthesis of beta-lactam antibiotics. Equilibrium concentrations and side-chain properties. Enzyme and Microbial Technology 24(8,Äì9): 498-506.

Shewale, J. G. and Sudhakaran, V. K. (1997). Penicillin V acylase: Its potential in the production of 6-aminopenicillanic acid. Enzyme and Microbial Technology 20(6): 402-410.

Shin, J., Ming, G.-l. and Song, H. (2014). Decoding neural transcriptomes and epigenomes via high-throughput sequencing. Nat Neurosci 17(11): 1463-1475.

Simon, C., Wiezer, A., Strittmatter, A. W. and Daniel, R. (2009). Phylogenetic Diversity and Metabolic Potential Revealed in a Glacier Ice Metagenome. Applied and Environmental Microbiology 75(23): 7519-7526.

Skatrud, P. L. and Queener, S. W. (1989). An electrophoretic molecular karyotype for an industrial strain of Cephalosporium acremonium. Gene 78(2): 331-338.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. and Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. Nature 321(6071): 674-679.

Srirangan, K., Orr, V., Akawi, L., Westbrook, A., Moo-Young, M. and Chou, C. P. (2013). Biotechnological advances on Penicillin G acylase: Pharmaceutical implications, unique expression mechanism and production strategies. Biotechnology Advances 31(8): 1319-1332.

Suzuki, S. and Hoa, P. T. P. (2012). Distribution of Quinolones, Sulfonamides, Tetracyclines in Aquatic Environment and Antibiotic Resistance in Indochina. Frontiers in Microbiology 3: 67.

Takahashi, T., Yamazaki, Y. and Kato, K. (1974). Substrate specificity of an α-amino acid ester hydrolase produced by Acetobacter turbidans A.T.C.C. 9325. Biochemical Journal 137(3): 497-503.

Takahashi, T., Yamazaki, Y., Kato, K. and Isono, M. (1972). Enzymic synthesis of cephalosporins. Journal of the American Chemical Society 94(11): 4035-4037.

Theilgaard, H., van Den Berg, M., Mulder, C., Bovenberg, R. and Nielsen, J. (2001). Quantitative analysis of Penicillium chrysogenum Wis54-1255 transformants overexpressing the penicillin biosynthetic genes. Biotechnol Bioeng 72(4): 379-388.

Thykaer, J. and Nielsen, J. (2003). Metabolic engineering of beta-lactam production. Metabolic Engineering 5(1): 56-69.

Tischer, W., Giesecke, U., Lang, G., RÖDer, A. and Wedekind, F. (1992). Biocatalytic 7-AminocephaIosporanic Acid Production. Annals of the New York Academy of Sciences 672(1): 502-509.

Torres-Cortés, G., Millán, V., Ramírez-Saad, H. C., Nisa-Martínez, R., Toro, N. and Martínez-Abarca, F. (2011). Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. Environmental Microbiology 13(4): 1101-1114.

Ullan, R. V., Campoy, S., Casqueiro, J., Fernandez, F. J. and Martin, J. F. (2007). Deacetylcephalosporin C production in Penicillium chrysogenum by expression of the isopenicillin N epimerization, ring expansion, and acetylation genes. Chem Biol 14(3): 329-339.

Unger, I. M., Goyne, K. W., Kennedy, A. C., Kremer, R. J., McLain, J. E. T. and Williams, C. F. (2013). Antibiotic Effects on Microbial Community Characteristics in Soils under Conservation Management Practices. Soil Sci. Soc. Am. J. 77(1): 100-112.

Valle, F., Balba´s, P., Merino, E. and Bollvar, F. (1991). The role of penicillin amidases in nature and in industry. Trends in Biochemical Sciences 16: 36-40.

van Langen, L. M., de Vroom, E., van Rantwijk, F. and Sheldon, R. Enzymatic synthesis of β-lactam antibiotics using penicillin-G acylase in frozen media. FEBS Letters 456(1): 89-92.

Veenstra, A. E., van Solingen, P., Bovenberg, R. A. L. and van der Voort, L. H. M. (1991). Strain improvement of Penicillium chrysogenum by recombinant DNA techniques. Journal of Biotechnology 17(1): 81-90.

Velasco, J., Luis Adrio, J., Angel Moreno, M., Diez, B., Soler, G. and Luis Barredo, J. (2000). Environmentally safe production of 7-aminodeacetoxycephalosporanic acid (7-ADCA) using recombinant strains of Acremonium chrysogenum. Nat Biotech 18(8): 857-861.

Větrovský, T. and Baldrian, P. (2013). Analysis of soil fungal communities by amplicon pyrosequencing: current approaches to data analysis and the introduction of the pipeline SEED. Biology and Fertility of Soils 49(8): 1027-1037.

Wall, D. H., Bardgett, R. D. and Kelly, E. (2010). Biodiversity in the dark. Nature Geosci 3(5): 297-298.

Wang, L., Ye, L.-J., Pan, Y. and Cao, Y. (2012). Two plate-based colorimetric assays for screening α-amino acid ester hydrolase with high synthesis/hydrolysis ratio. Enzyme and Microbial Technology 51(2): 107-112.

Wang, Y., Zhang, Z. and Ramanan, N. (1997). The actinomycete Thermobispora bispora contains two distinct types of transcriptionally active 16S rRNA genes. Journal of Bacteriology 179(10): 3270-3276.

Wang, Z., Zhang, X.-X., Huang, K., Miao, Y., Shi, P., Liu, B., Long, C. and Li, A. (2013). Metagenomic Profiling of Antibiotic Resistance Genes and Mobile Genetic Elements in a Tannery Wastewater Treatment Plant. PLoS ONE 8(10): e76079.

Ward, J. M. and Hodgson, J. E. (1993). The biosynthetic genes for clavulanic acid and cephamycin production occur as a "super-cluster" in three Streptomyces. FEMS Microbiology Letters 110(2): 239-242.

Weerasinghe, C. A. and Towner, D. (1997). Aerobic biodegradation of virginiamycin in soil. Environmental Toxicology and Chemistry 16(9): 1873-1876.

Wegman, Margreth A., Janssen, Michiel H. A., van Rantwijk, F. and Sheldon, Roger A. (2001). Towards Biocatalytic Synthesis of β-Lactam Antibiotics. Advanced Synthesis & Catalysis 343(6-7): 559-576.

Weil, J., Miramonti, J. and Ladisch, M. R. (1995). Cephalosporin C: Mode of action and biosynthetic pathway. Enzyme and Microbial Technology 17(1): 85-87.

Werner, J. J., Zhou, D., Caporaso, J. G., Knight, R. and Angenent, L. T. (2012). Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J 6(7): 1273-1276.

Wu, Q., Chen, C.-X., Du, L.-L. and Lin, X.-F. (2010). Enzymatic Synthesis of Amoxicillin via a One-pot Enzymatic Hydrolysis and Condensation Cascade Process in the Presence of Organic Co-solvents. Applied Biochemistry and Biotechnology 160(7): 2026-2035.

Xue, B., Zhang, R., Wang, Y., Liu, X., Li, J. and Zhang, G. (2013). Antibiotic contamination in a typical developing city in south China: Occurrence and ecological risks in the Yongjiang River impacted by tributary discharge and anthropogenic activities. Ecotoxicology and Environmental Safety 92(0): 229-236.

Yang, X.-M. (1992). Optimization of a cultivation process for recombinant protein production by Escherichia coli. Journal of Biotechnology 23(3): 271-289.

Yang, Y., Biedendieck, R., Wang, W., Gamer, M., Malten, M., Jahn, D. and Deckwer, W. D. (2006). High yield recombinant penicillin G amidase production and export into the growth medium using Bacillus megaterium. Microb Cell Fact 5: 36.

Ye, L.-J., Wang, L., Pan, Y. and Cao, Y. (2012). Changing the specificity of α-amino acid ester hydrolase toward para-hydroxyl cephalosporins synthesis by site-directed saturation mutagenesis. Biotechnology Letters 34(9): 1719-1724.

Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A. and Elshahed, M. S. (2009). Comparison of Species Richness Estimates Obtained Using Nearly Complete Fragments and Simulated Pyrosequencing-Generated Fragments in 16S rRNA Gene-Based Environmental Surveys. Applied and Environmental Microbiology 75(16): 5227-5236.

Zarubina, S. A., Uporov, I. V., Fedorchuk, E. A., Fedorchuk, V. V., Sklyarenko, A. V., Yarotsky, S. V. and Tishkov, V. I. (2013). 3D Structure Modeling of Alpha-Amino Acid Ester Hydrolase from Xanthomonas rubrilineans. Acta Naturae 5(4): 62-70.

Zhalnina, K., Dias, R., de Quadros, P., Davis-Richardson, A., Camargo, F. O., Clark, I., McGrath, S., Hirsch, P. and Triplett, E. (2015). Soil pH Determines Microbial Diversity and Composition in the Park Grass Experiment. Microbial Ecology 69(2): 395-406.

Zhan, Y., Yan, Y., Zhang, W., Chen, M., Lu, W., Ping, S. and Lin, M. (2012). Comparative analysis of the complete genome of an Acinetobacter calcoaceticus strain adapted to a phenol-polluted environment. Research in Microbiology 163(1): 36-43.

Zhang, M., Shi, M., Zhou, Z., Yang, S., Yuan, Z. and Ye, Q. (2006). Production of Alcaligenes faecalis penicillin G acylase in Bacillus subtilis WB600 (pMA5) fed with partially hydrolyzed starch. Enzyme and Microbial Technology 39(4): 555-560.

Zhou, J., Xia, B., Huang, H., Treves, D. S., Hauser, L. J., Mural, R. J., Palumbo, A. V. and Tiedje, J. M. (2003). Bacterial phylogenetic diversity and a novel candidate division of two humid region, sandy surface soils. Soil Biology and Biochemistry 35(7): 915-924.